

REFERENCE/DOCKET NUMBER: 210121.440C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
US-09-073-010-73

Query Match 2.8%; Score 6; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 SDVLTA 188
Db 8 SDVLTA 13

RESULT 22
US-09-073-010-75
Sequence 75, Application US/09073010
Patent No. 6613881
GENERAL INFORMATION:
APPLICANT: Alderson, Mark
APPLICANT: Dillon, David C.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Campos-Neco, Antonio
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS OF TUBERCULOSIS AND
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY
STREET: 6300 Columbia Center, 701 Fifth Ave.
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,010
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.440C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
US-09-073-010-75

Query Match 2.8%; Score 6; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 SDVLTA 188
Db 3 SDVLTA 8

RESULT 23
US-07-895-252-10
Sequence 10, Application US/07895252
Patent No. 5278136
GENERAL INFORMATION:
APPLICANT: Skubitz, Amy P.N.
APPLICANT: Furcht, Leo T.
TITLE OF INVENTION: LAMININ A CHAIN POLYPEPTIDES FROM
THE AMINO TERMINAL GLOBULAR DOMAIN
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant & Gould
STREET: 3100 No. 5276136west Center
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/895,252
FILING DATE: 19920608
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/646,291
FILING DATE: 25-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kowalchuk, Alan W.
REGISTRATION NUMBER: 31,535
REFERENCE/DOCKET NUMBER: 600.212-US-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-332-5300
TELEFAX: 612-332-9081
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Synthetically derived
US-07-895-252-10

Query Match 2.8%; Score 6; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSAD 18
Db 12 VAVSAD 17

RESULT 24
US-08-072-283B-10
Sequence 10, Application US/08072283B
Patent No. 5703205
GENERAL INFORMATION:
APPLICANT: Skubitz, Amy P.N.
APPLICANT: Furcht, Leo T.
TITLE OF INVENTION: LAMININ A CHAIN POLYPEPTIDES
FROM THE AMINO TERMINAL GLOBULAR DOMAIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt

STREET: 3100 No. 5703205 West Center, 90 South Seventh St
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/072,283B
FILING DATE: 07-JUN-1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/895,252
FILING DATE: 08-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 600.212-US-02
TELECOMMUNICATION INFORMATION:
TELEPHONE:
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5703205e
US-08-072-283B-10

Query Match 2.8%; Score 6; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSAD 18
DB 12 VAVSAD 17

RESULT 25
US-07-998-820-2
Sequence 2, Application US/07998820
Patent No. 5738838
GENERAL INFORMATION:
APPLICANT: Zamora, Paul O.
TITLE OF INVENTION: IKVAV Peptide
TITLE OF INVENTION: Radiopharmaceutical Applications
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: RhoMed Incorporated
STREET: 4261 Balloon Park
CITY: Albuquerque
STATE: NM
COUNTRY: U.S.A.
ZIP: 87109-5802
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb
MEDIUM TYPE: Storage
COMPUTER: IBM PC/XT/AT, IBM PS/2 or
COMPUTER: compatibles
OPERATING SYSTEM: PC-DOS or MS-DOS
SOFTWARE: WordPerfect 6.0a for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/998,820
FILING DATE: 30-DEC-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/840,077
FILING DATE: 20-FEB-1992

ATTORNEY/AGENT INFORMATION:
NAME: David P. Hegge
REGISTRATION NUMBER: 36,827
TELECOMMUNICATION INFORMATION:
TELEPHONE: (505) 344-7200
TELEFAX: (505) 344-9460
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-07-998-820-2

Query Match 2.8%; Score 6; DB 1; Length 19;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSAD 18
DB 13 VAVSAD 18

RESULT 26
US-08-280-646-3
Sequence 3, Application US/08280646
Patent No. 5834029
GENERAL INFORMATION:
APPLICANT: Bellamkonda, Ravi
APPLICANT: Ranieri, John P
APPLICANT: Aebischer, Patrick
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR A
TITLE OF INVENTION: BICARTIFICIAL EXTRACELLULAR MATRIX
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: James F. Haley, Jr., FISH & NEAVE
STREET: 1251 Ave. of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/280,646
FILING DATE: 20-JUL-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: CTI-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-280-646-3

Query Match 2.8%; Score 6; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 39;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSAD 18
| | | | |
Db 13 VAVSAD 18

RESULT 27
US-08-445-193-3
; Sequence 3, Application US/08445193
; Patent No. 5840576
; GENERAL INFORMATION:
; APPLICANT: Schinstine, Malcolm
; APPLICANT: Shoichet, Molly S.
; APPLICANT: Gentile, Frank T.
; APPLICANT: Hamang, Joseph P.
; APPLICANT: Holland, Laura M.
; APPLICANT: Cain, Brian M.
; APPLICANT: Doherty, Edward J.
; APPLICANT: Winn, Shelley R.
; APPLICANT: Aebischer, Patrick
; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF GROWTH CONTROL
; TITLE OF INVENTION: FOR CELLS ENCAPSULATED IN BIOARTIFICIAL ORGANS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., FISH & NEAVE
; STREET: 1251 Ave. of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,193
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; FILING DATE: 20-JULY-1994
; REGISTRATION NUMBER: 27,794
; APPLICATION NUMBER: US 08/279,773
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REFERENCE/DOCKET NUMBER: CTI-22 CIP
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-445-193-3

Query Match 2.8%; Score 6; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSAD 18
| | | | |
Db 13 VAVSAD 18

RESULT 28

US-08-432-698-3
; Sequence 3, Application US/08432698
; Patent No. 5843431
; GENERAL INFORMATION:
; APPLICANT: Schinstine, Malcolm
; APPLICANT: Shoichet, Molly S.
; APPLICANT: Gentile, Frank T.
; APPLICANT: Hamang, Joseph P.
; APPLICANT: Holland, Laura M.
; APPLICANT: Cain, Brian M.
; APPLICANT: Doherty, Edward J.
; APPLICANT: Winn, Shelley R.
; APPLICANT: Aebischer, Patrick
; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF GROWTH CONTROL
; TITLE OF INVENTION: FOR CELLS ENCAPSULATED IN BIOARTIFICIAL ORGANS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., FISH & NEAVE
; STREET: 1251 Ave. of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,698
; FILING DATE: 09-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; FILING DATE: 20-JULY-1994
; REGISTRATION NUMBER: 27,794
; APPLICATION NUMBER: CTI-22 CIP
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-432-698-3

Query Match 2.8%; Score 6; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSAD 18
| | | | |
Db 13 VAVSAD 18

RESULT 29
US-08-447-810-3
; Sequence 3, Application US/08447810
; Patent No. 5858747
; GENERAL INFORMATION:
; APPLICANT: Schinstine, Malcolm
; APPLICANT: Shoichet, Molly S.
; APPLICANT: Gentile, Frank T.
; APPLICANT: Hamang, Joseph P.
; APPLICANT: Holland, Laura M.
; APPLICANT: Cain, Brian M.

APPLICANT: Doherty, Edward J.
APPLICANT: Winn, Shelley R.
APPLICANT: Aebischer, Patrick
TITLE OF INVENTION: METHODS AND COMPOSITIONS OF GROWTH CONTROL
FOR CELLS ENCAPSULATED IN BIOARTIFICIAL ORGANS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: James F. Haley, Jr., FISH & NEAVE
STREET: 1251 Ave. of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,810
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/432,698
FILING DATE: 09-MAY-1995
APPLICATION NUMBER: US 08/279,773
FILING DATE: 20-JULY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: CTI-22 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-447-810-3

Query Match 2.8%; Score 6; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 VAVSAD 18
Db 13 VAVSAD 18

RESULT 30
US-08-279-773-3
Sequence 3, Application US/08279773
Patent No. 5935849
GENERAL INFORMATION:
APPLICANT: Schinstine, Malcolm
APPLICANT: Shoichet, Molly S.
APPLICANT: Gentile, Frank T.
APPLICANT: Hamming, Joseph P.
APPLICANT: Holland, Laura M.
APPLICANT: Cain, Brian M.
APPLICANT: Doherty, Edward J.
APPLICANT: Winn, Shelley R.
APPLICANT: Aebischer, Patrick
TITLE OF INVENTION: METHODS AND COMPOSITIONS OF GROWTH CONTROL
FOR CELLS ENCAPSULATED IN BIOARTIFICIAL ORGANS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: James F. Haley, Jr., FISH & NEAVE

STREET: 1251 Ave. of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/279,773
FILING DATE: 20-JUL-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: CTI-22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-279-773-3

Query Match 2.8%; Score 6; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 VAVSAD 18
Db 13 VAVSAD 18

RESULT 31
US-08-747-137-56
Sequence 56, Application US/08747137
Patent No. 5945033
GENERAL INFORMATION:
APPLICANT: Yen, Richard C.K.
TITLE OF INVENTION: NON-CROSSLINKED PROTEIN PARTICLES FOR
THERAPEUTIC AND DIAGNOSTIC USE
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/747,137
FILING DATE: 12-NOV-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,546
FILING DATE: 14-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/069,831
FILING DATE: 01-JUN-1993
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/959,560
FILING DATE: 13-OCT-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/641,720
FILING DATE: 15-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 016197-000840US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
US-08-747-137-56

Query Match 2.8%; Score 6; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSAD 18
Db 13 VAVSAD 18

RESULT 32
US-08-447-997-3
Sequence 3, Application US/08447997
Patent No. 6392118
GENERAL INFORMATION:
APPLICANT: Schinstine, Malcolm
APPLICANT: Shoichet, Molly S.
APPLICANT: Gentile, Frank T.
APPLICANT: Hamman, Joseph P.
APPLICANT: Holland, Laura M.
APPLICANT: Cain, Brian M.
APPLICANT: Doherty, Edward J.
APPLICANT: Winn, Shelley R.
APPLICANT: Aebischer, Patrick
TITLE OF INVENTION: METHODS AND COMPOSITIONS OF GROWTH CONTROL
TITLE OF INVENTION: FOR CELLS ENCAPSULATED IN BIOARTIFICIAL ORGANS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: James F. Haley, Jr., FISH & NEAVE
STREET: 1251 Ave. of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020-1104
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,997
FILING DATE: 23-MAY-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/432,698
FILING DATE: 09-MAY-1995
APPLICATION NUMBER: US 08/279,773
FILING DATE: 20-JULY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: CFI-22 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-447-997-3

Query Match 2.8%; Score 6; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSAD 18
Db 13 VAVSAD 18

RESULT 33
US-09-801-237A-3
Sequence 3, Application US/09801237A
Patent No. 6495364
GENERAL INFORMATION:
APPLICANT: Hamman et al.
TITLE OF INVENTION: MX-1 Conditionally Immortalized Cells
FILE REFERENCE: 19141-534 CON
CURRENT APPLICATION NUMBER: US/09/801,237A
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: 08/279,773
PRIOR FILING DATE: 1994-07-20
PRIOR APPLICATION NUMBER: 08/447,997
PRIOR FILING DATE: 1995-05-23
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 3
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:chemically
OTHER INFORMATION: synthesized
US-09-801-237A-3

Query Match 2.8%; Score 6; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSAD 18
Db 13 VAVSAD 18

RESULT 34
PCT-US95-09281-3
Sequence 3, Application PC/TUS9509281
GENERAL INFORMATION:
APPLICANT: Cytotherapeutics, Inc.
TITLE OF INVENTION: METHODS AND COMPOSITIONS OF GROWTH CONTROL
TITLE OF INVENTION: FOR CELLS ENCAPSULATED IN BIOARTIFICIAL ORGANS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: James F. Haley, Jr., FISH & NEAVE
STREET: 1251 Ave. of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020-1104
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,997
FILING DATE: 23-MAY-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/432,698
FILING DATE: 09-MAY-1995
APPLICATION NUMBER: US 08/279,773
FILING DATE: 20-JULY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: CFI-22 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090

;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/09281
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/279,773
;; FILING DATE: 20-JULY-1994
;; APPLICATION NUMBER: US 08/432,698
;; FILING DATE: 09-MAY-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Haley Jr., James F.
;; REGISTRATION NUMBER: 27,794
;; REFERENCE/DOCKET NUMBER: CTI-22 CIP PCT
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 596-9000
;; TELEFAX: (212) 596-9090
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 19 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
PCT-US95-09281-3

Query Match 2.8%; Score 6; DB 5; Length 19;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

QY 13 VAVSAD 18
Db 13 VAVSAD 18

RESULT 35
PCT-US95-09282-3
;; Sequence 3, Application PC/TUS9509282
;; GENERAL INFORMATION:
;; APPLICANT: Cytotherapeutics, Inc.
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR A
;; TITLE OF INVENTION: BIOARTIFICIAL EXTRACELLULAR MATRIX
;; NUMBER OF SEQUENCES: 4
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: James F. Haley, Jr., FISH & NEAVE
;; STREET: 1251 Ave. of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 10020-1104
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/09282
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/280,646
;; FILING DATE: 20-JULY-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Haley Jr., James F.
;; REGISTRATION NUMBER: 27,794
;; REFERENCE/DOCKET NUMBER: CTI-26 PCT
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 596-9000
;; TELEFAX: (212) 596-9090
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 19 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
PCT-US95-09282-3

Query Match 2.8%; Score 6; DB 5; Length 19;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

QY 13 VAVSAD 18
Db 13 VAVSAD 18

RESULT 36
US-09-562-737-124
;; Sequence 124, Application US/09562737
;; Patent No. 6428967
;; GENERAL INFORMATION:
;; APPLICANT: Herz, Joachim
;; TITLE OF INVENTION: LDL Receptor Signaling Pathways
;; FILE REFERENCE: UTSW0708
;; CURRENT APPLICATION NUMBER: US/09/562,737
;; CURRENT FILING DATE: 2000-05-01
;; NUMBER OF SEQ ID NOS: 132
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 124
;; LENGTH: 57
;; TYPE: PRT
;; ORGANISM: mouse
US-09-562-737-124

Query Match 2.8%; Score 6; DB 4; Length 57;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

QY 140 VALSLE 145
Db 47 VALSLE 52

RESULT 37
US-08-905-223-335
;; Sequence 335, Application US/08905223
;; Patent No. 6222029
;; GENERAL INFORMATION:
;; APPLICANT: Edwards, Jean-Baptiste D.
;; APPLICANT: Duclert, Aymeric
;; APPLICANT: Lacroix, Bruno
;; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
;; NUMBER OF SEQUENCES: 503
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Kucbbe, Martens, Olson & Bear
;; STREET: 501 West Broadway
;; CITY: San Diego
;; STATE: California
;; COUNTRY: USA
;; ZIP: 92101-3505
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy Disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: Win95
;; SOFTWARE: Word
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/905,223
;; FILING DATE:
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 335:
SEQUENCE CHARACTERISTICS:
LENGTH: 67 amino acids
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: sig_peptide
LOCATION: -43...-1
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 5.2
OTHER INFORMATION: seq PWQLQLVLSVDVLA/EI
US-08-905-223-335

Query Match 2.8%; Score 6; DB 3; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 VLSDVL 186
DB 37 VLSDVL 42

RESULT 38

US-09-107-532A-6800
Sequence 6800, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Arianello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-5007
TELEFAX: (781) 893-8277
INFORMATION FOR SEQ ID NO: 6800:
SEQUENCE CHARACTERISTICS:
LENGTH: 67 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...67
SEQUENCE DESCRIPTION: SEQ ID NO: 6800:
US-09-107-532A-6800

Query Match 2.8%; Score 6; DB 4; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 HLLIGV 107
DB 46 HLLIGV 51

RESULT 39

US-09-543-681A-5866
Sequence 5866, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:

APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/129,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 5866
LENGTH: 71
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-5866

Query Match 2.8%; Score 6; DB 4; Length 71;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 186 LTAIFQ 191
DB 18 LTAIFQ 23

RESULT 40

US-09-634-238-234
Sequence 234, Application US/09634238
Patent No. 6544772
GENERAL INFORMATION:

APPLICANT: Glenn, Matthew
APPLICANT: Havukkala, Ilkka J.
APPLICANT: Bloksberg, Leonard, N.
APPLICANT: Lubbers, Mark W.
APPLICANT: Dekker, James
APPLICANT: Christensson, Anna C.
APPLICANT: Holland, Ross
APPLICANT: O'Toole, Paul W.
APPLICANT: Reid, Julian R.
APPLICANT: Coolbear, Timothy
TITLE OF INVENTION: Polynucleotides, materials incorporating
them and methods for using them.

FILE REFERENCE: 11000.1043U1
CURRENT APPLICATION NUMBER: US/09/634,238
CURRENT FILING DATE: 2000-08-08
NUMBER OF SEQ ID NOS: 422

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 234
LENGTH: 88
TYPE: PRT

; ORGANISM: Lactobacillus rhamnosus
US-09-634-238-234

Query Match 2.8%; Score 6; DB 4; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 35 DAIAAI 40
Db 72 DAIAAI 77

RESULT 41

US-09-543-681A-6613
; Sequence 6613, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETTON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; PRIOR FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 6613

; LENGTH: 91

; TYPE: PRT

; ORGANISM: Proteus mirabilis

US-09-543-681A-6613

Query Match

Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 33 IDDAIA 38
Db 22 IDDAIA 27

RESULT 42

US-09-073-009-25

; Sequence 25, Application US/09073009

; Patent No. 6555653

; GENERAL INFORMATION:

; APPLICANT: Alderson, Mark

; APPLICANT: Dillon, Davin C.

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Campos-Neto, Antonio

; TITLE OF INVENTION: COMPOUNDS AND DIAGNOSIS OF

; NUMBER OF SEQUENCES: 144

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY

; STREET: 6300 Columbia Center, 701 Fifth Ave.

; CITY: Seattle

; STATE: Washington

; COUNTRY: US

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/073,009

; FILING DATE: 05-MAY-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Makl, David J.

; REGISTRATION NUMBER: 31,392

; REFERENCE/DOCKET NUMBER: 210121.441C1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 206-622-4900

; TELEFAX: 206-682-6031

; INFORMATION FOR SEQ ID NO: 25:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 94 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; ORIGINAL SOURCE: Mycobacterium tuberculosis

US-09-073-009-25

Query Match

Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 183 SDVLTA 188
Db 33 SDVLTA 38

RESULT 43

US-09-073-009-27

; Sequence 27, Application US/09073009

; Patent No. 6555653

; GENERAL INFORMATION:

; APPLICANT: Alderson, Mark

; APPLICANT: Dillon, Davin C.

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Campos-Neto, Antonio

; TITLE OF INVENTION: COMPOUNDS AND DIAGNOSIS OF

; NUMBER OF SEQUENCES: 144

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY

; STREET: 6300 Columbia Center, 701 Fifth Ave.

; CITY: Seattle

; STATE: Washington

; COUNTRY: US

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/073,009

; FILING DATE: 05-MAY-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Makl, David J.

; REGISTRATION NUMBER: 31,392

; REFERENCE/DOCKET NUMBER: 210121.441C1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 206-622-4900

; TELEFAX: 206-682-6031

; INFORMATION FOR SEQ ID NO: 27:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 94 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; ORIGINAL SOURCE: Mycobacterium tuberculosis

US-09-073-009-27

Query Match

Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 183 SDVLTA 188
Db 33 SDVLTA 38

TUBERCULOSIS AND MET

TUBERCULOSIS AND ME

Db 33 SDVLTA 38

RESULT 44

US-09-073-009-29

; Sequence 29, Application US/09073009

; Patent No. 655563

; GENERAL INFORMATION:

; APPLICANT: Alderson, Mark

; APPLICANT: Dillon, Davin C.

; APPLICANT: Skeiky, Yasir A.W.

; TITLE OF INVENTION: COMPOUNDS AND DIAGNOSIS OF

; NUMBER OF SEQUENCES: 144

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY

; STREET: 6300 Coubia Center, 701 Fifth Ave.

; City: Seattle

; STATE: Washington

; COUNTRY: US

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/073,009

; FILING DATE: 05-MAY-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Maki, David J.

; REGISTRATION NUMBER: 31,392

; REFERENCE/DOCKET NUMBER: 210121.441C1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 206-622-4900

; TELEFAX: 206-682-6031

; INFORMATION FOR SEQ ID NO: 29:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 94 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; ORIGINAL SOURCE:

; ORGANISM: Mycobacterium tuberculosis

; US-09-073-009-29

Query Match 2.8%; Score 6; DB 4; Length 94;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 SDVLTA 188

|||||

Db 33 SDVLTA 38

RESULT 45

US-09-073-010-25

; Sequence 25, Application US/09073010

; Patent No. 6613881

; GENERAL INFORMATION:

; APPLICANT: Alderson, Mark

; APPLICANT: Dillon, Davin C.

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Campos-Neto, Antonio

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS OF TUBERCULOSIS AND

; NUMBER OF SEQUENCES: 144

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY

; STREET: 6300 Coubia Center, 701 Fifth Ave.

; City: Seattle

; STATE: Washington

COUNTRY: US

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/073,010

FILING DATE: 05-MAY-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.440C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-622-4900

TELEFAX: 206-682-6031

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 94 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: Mycobacterium tuberculosis

US-09-073-010-25

Query Match

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 SDVLTA 188

|||||

Db 33 SDVLTA 38

RESULT 46

US-09-073-010-27

; Sequence 27, Application US/09073010

; Patent No. 6613881

; GENERAL INFORMATION:

; APPLICANT: Alderson, Mark

; APPLICANT: Dillon, Davin C.

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Campos-Neto, Antonio

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS OF TUBERCULOSIS AND

; NUMBER OF SEQUENCES: 144

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY

; STREET: 6300 Coubia Center, 701 Fifth Ave.

; City: Seattle

; STATE: Washington

; COUNTRY: US

; ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/073,010

FILING DATE: 05-MAY-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.440C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-622-4900

TELEFAX: 206-682-6031

INFORMATION FOR SEQ ID NO: 27:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 94 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Mycobacterium tuberculosis
US-09-073-010-27

Query Match 2.8%; Score 6; DB 4; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 SDVLTA 188
Db 33 SDVLTA 38

RESULT 47
US-09-073-010-29
; Sequence 29, Application US/09073010
; Patent No. 6613881
; GENERAL INFORMATION:
; APPLICANT: Alderson, Mark
; APPLICANT: Dillion, David C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Campos-Neto, Antonio
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS OF TUBERCULOSIS AND
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY
; STREET: 6300 Columbia Center, 701 Fifth Ave.
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,010
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.440C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; TELEFAX: 206-682-6031
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 94 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Mycobacterium tuberculosis
US-09-073-010-29

Query Match 2.8%; Score 6; DB 4; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 SDVLTA 188
Db 33 SDVLTA 38

RESULT 48
US-09-252-991A-24298
; Sequence 24298, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24298
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24298

Query Match 2.8%; Score 6; DB 4; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 ARGKQ 83
Db 37 ARGKQ 42

RESULT 49
US-09-383-586-34
; Sequence 34, Application US/09383586
; Patent No. 6242419
; GENERAL INFORMATION:
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Abernethy, Nevin
; APPLICANT: Onrust, Rene
; APPLICANT: Kumble, Anand
; APPLICANT: Murison, Greg
; TITLE OF INVENTION: Compounds isolated from stromal cells
; FILE REFERENCE: 11000.1037c1
; CURRENT APPLICATION NUMBER: US/09/383,586
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Mouse
US-09-383-586-34

Query Match 2.8%; Score 6; DB 3; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MKFLLI 7
Db 1 MKFLLI 6

RESULT 50
US-09-540-236-2114
; Sequence 2114, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRH
; FILE REFERENCE: 2709.2005-001

```
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2114
; LENGTH: 113
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-2114

Query Match          2.8%; Score 6; DB 4; Length 113;
Best Local Similarity 100.0%; Pred. No. 2.e+02; 0; Indels 0;
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 41 EQSETI 46
Db 73 EQSETI 78

RESULT 51
5164490-6
; Patent No. 5164490
; APPLICANT: SANTI, DANIEL V.; EDMAN, JEFFREY; EDMAN, URSULA
; TITLE OF INVENTION: PNEUMOCYSTIS CARINII DIHYDROFOLATE
; REDUCTASE GENE AND METHODS FOR ITS USE
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/437,511
; FILING DATE: 15-NOV-1989
; SEQ ID NO: 6;
; LENGTH: 115
5164490-6

Query Match          2.8%; Score 6; DB 6; Length 115;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

QY 180 GVLSDV 185
Db 93 GVLSDV 98

RESULT 52
US-09-562-737-125
; Sequence 125, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 125
; LENGTH: 116
; TYPE: PRT
; ORGANISM: mouse
US-09-562-737-125

Query Match          2.8%; Score 6; DB 4; Length 116;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

QY 140 VALSLE 145
Db 106 VALSLE 111

RESULT 53
US-09-199-637A-150
; Sequence 150, Application US/09199637A
; Patent No. 6355411
```

```
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-150

Query Match          2.8%; Score 6; DB 4; Length 128;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

QY 170 GGLSIL 175
Db 23 GGLSIL 28

RESULT 54
US-08-944-449-2
; Sequence 2, Application US/08944449
; Patent No. 5985613
; GENERAL INFORMATION:
; APPLICANT: KURTH, REINHARD
; APPLICANT: BAIER, MICHAEL
; APPLICANT: METZNER, KARIN
; APPLICANT: WERNER, ALBRECHT
; TITLE OF INVENTION: Use of an "immunodeficiency-virus suppressing
; TITLE OF INVENTION: lymphokine (ISL)" to inhibit the replication of
; TITLE OF INVENTION: viruses, particularly of retroviruses
; FILE REFERENCE: 8341-7065
; CURRENT APPLICATION NUMBER: US/08/944,449
; CURRENT FILING DATE: 1997-10-06
; EARLIER APPLICATION NUMBER: EP 95113013.2
; EARLIER FILING DATE: 1995-08-18
; EARLIER APPLICATION NUMBER: DE 195 13 152.5
; EARLIER FILING DATE: 1995-04-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 130
; TYPE: PRT
; ORGANISM: african green monkey
US-08-944-449-2

Query Match          2.8%; Score 6; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

QY 41 EQSETI 46
Db 71 EQSETI 76

RESULT 55
US-09-353-362-2
; Sequence 2, Application US/09353362
; Patent No. 6383739
```

```

; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Use of an "immunodeficiency-virus suppressing
; LYMPHOKINE (ISL)" to inhibit the replication of viruses,
; TITLE OF INVENTION: in particular of retroviruses
; NUMBER OF SEQUENCES: 8
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30B (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/353,362
; FILING DATE: 15-JUL-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 13 152.5
; FILING DATE: 07-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95113013.2
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: KLESNER, Sharon N.
; REGISTRATION NUMBER: 36,335
; REFERENCE/DOCKET NUMBER: P8341-9012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-5000
; TELEFAX: (202) 638-4810
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 130 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-353-362-2

Query Match 2.8%; Score 6; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 EQSETI 46
Db 71 EQSETI 76

RESULT 56
US-08-891-848-15
; Sequence 15, Application US/08891848
; Patent No. 5955073
; GENERAL INFORMATION:
; APPLICANT: Rybak, Susanna M.
; APPLICANT: Foule, Richard J.
; APPLICANT: Newton, Dianne L.
; APPLICANT: Nicholls, Peter J.
; TITLE OF INVENTION: Selective RNase Cytotoxic Reagents
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,848
; FILING DATE: No. 5955073 yet assigned
; CLASSIFICATION: 530

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/125,462
; FILING DATE: 22-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/014,082
; FILING DATE: 04-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,195
; FILING DATE: 22-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/510,696
; FILING DATE: 20-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Leaver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 015280-110310US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..133
; OTHER INFORMATION: /note="Human eosinophil cationic
; protein (ECP)"
; US-08-891-848-15

Query Match 2.8%; Score 6; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 162 FANVYN 167
Db 48 FANVYN 53

RESULT 57
US-08-875-811-10
; Sequence 10, Application US/08875811
; Patent No. 6045793
; GENERAL INFORMATION:
; APPLICANT: Rybak, Susanna M.
; APPLICANT: Newton, Dianne L.
; APPLICANT: Boque, Lluís
; APPLICANT: Wlodawer, Alexander
; TITLE OF INVENTION: Recombinant Ribonuclease Proteins
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,811
; FILING DATE: 19-FEB-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/02588
; FILING DATE: 19-FEB-1997
; PRIOR APPLICATION DATA:

```


/ APPLICATION NUMBER: US 60/011,800
/ FILING DATE: 21-FEB-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Faris, Susan K.
/ REGISTRATION NUMBER: 41,739
/ REFERENCE/DOCKET NUMBER: 015280-244100US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0200
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 133 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ FEATURE:
/ NAME/KEY: Protein
/ LOCATION: 1..133
/ OTHER INFORMATION: /note= "Human eosinophil cationic
/ OTHER INFORMATION: Protein (ECP)"
US-08-875-811-10

Query Match 2.8%; Score 6; DB 3; Length 133;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 162 FANVVN 167
Db 48 FANVVN 53

RESULT 58
US-08-891-848-14
/ Sequence 14, Application US/08891848
/ Patent No. 5955073
/ GENERAL INFORMATION:
/ APPLICANT: Rybak, Susanna M.
/ APPLICANT: Youle, Richard J.
/ APPLICANT: Newton, Dianne L.
/ APPLICANT: Nicholls, Peter J.
/ TITLE OF INVENTION: Selective RNase Cytotoxic Reagents
/ NUMBER OF SEQUENCES: 19
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Townsend and Townsend and Crew LLP
/ STREET: Two Embarcadero Center, Eighth Floor
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94111-3834
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/891,848
/ FILING DATE: No. 5955073 yet assigned
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/125,462
/ FILING DATE: 22-SEP-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/014,082
/ FILING DATE: 04-FEB-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/779,195
/ FILING DATE: 22-OCT-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/510,696
/ FILING DATE: 20-APR-1990
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Weber, Ellen Lauver

/ REGISTRATION NUMBER: 32,762
/ REFERENCE/DOCKET NUMBER: 015280-110310US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0200
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 134 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ FEATURE:
/ NAME/KEY: Protein
/ LOCATION: 1..134
/ OTHER INFORMATION: /note= "Human eosinophil derived
/ OTHER INFORMATION: neurotoxin (EDN)"
US-08-891-848-14

Query Match 2.8%; Score 6; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 162 FANVVN 167
Db 48 FANVVN 53

RESULT 59
US-08-875-811-9
/ Sequence 9, Application US/08875811
/ Patent No. 6045793
/ GENERAL INFORMATION:
/ APPLICANT: Rybak, Susanna M.
/ APPLICANT: Newton, Dianne L.
/ APPLICANT: Boque, Lluís
/ TITLE OF INVENTION: Recombinant Ribonuclease Proteins
/ NUMBER OF SEQUENCES: 64
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Townsend and Townsend and Crew LLP
/ STREET: Two Embarcadero Center, Eighth Floor
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94111-3834
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/875,811
/ FILING DATE: 19-FEB-1998
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: WO PCT/US97/02588
/ FILING DATE: 19-FEB-1997
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 60/011,800
/ FILING DATE: 21-FEB-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Faris, Susan K.
/ REGISTRATION NUMBER: 41,739
/ REFERENCE/DOCKET NUMBER: 015280-244100US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0200
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 134 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:

TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..134
OTHER INFORMATION: /note= "Human eosinophil-derived
OTHER INFORMATION: neurotoxin (EDN)"
US-08-875-811-9

Query Match 2.8%; Score 6; DB 3; Length 134;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 FANVNV 167
Db 48 FANVNV 53

RESULT 60
US-09-446-959-5
; Sequence 5, Application US/09446959
; Patent No. 6426070
; GENERAL INFORMATION:
; APPLICANT: ROSENBERG, HELENE F.
; APPLICANT: DOMACHOWSKIE, JOSEPH B.
; TITLE OF INVENTION: METHODS FOR INACTIVATING ENVELOPED RNA VIRUS PARTICLES
; FILE REFERENCE: 11613.8USWO
; CURRENT APPLICATION NUMBER: US/09/446,959
; CURRENT FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: PCT/US98/13852
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/052,986
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-446-959-5

Query Match 2.8%; Score 6; DB 4; Length 134;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 FANVNV 167
Db 48 FANVNV 53

RESULT 61
US-09-252-991A-25498
; Sequence 25498, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25498
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25498

Query Match 2.8%; Score 6; DB 4; Length 136;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 147 SDEGNI 152
Db 86 SDEGNI 91

RESULT 62
US-09-383-586-35
; Sequence 35, Application US/09383586
; Patent No. 6242419
; GENERAL INFORMATION:
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Abernethy, Nevin
; APPLICANT: Onrust, Rene
; APPLICANT: Kumble, Anand
; APPLICANT: Muriison, Greg
; TITLE OF INVENTION: Compounds isolated from stromal cells
; FILE REFERENCE: 11000.1037c1
; CURRENT APPLICATION NUMBER: US/09/383,586
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Mouse
US-09-383-586-35

Query Match 2.8%; Score 6; DB 3; Length 147;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MKFLLI 7
Db 1 MKFLLI 6

RESULT 63
US-09-446-959-7
; Sequence 7, Application US/09446959
; Patent No. 6426070
; GENERAL INFORMATION:
; APPLICANT: ROSENBERG, HELENE F.
; APPLICANT: DOMACHOWSKIE, JOSEPH B.

US-09-446-959-7
; TITLE OF INVENTION: METHODS FOR INACTIVATING ENVELOPED RNA VIRUS PARTICLES
; FILE REFERENCE: 11613.8USWO
; CURRENT APPLICATION NUMBER: US/09/446,959
; CURRENT FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: PCT/US98/13852
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/052,986
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-446-959-7

Query Match 2.8%; Score 6; DB 4; Length 152;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 FANVNV 167
Db 49 FANVNV 54

RESULT 64
US-09-489-039A-11904
; Sequence 11904, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11904
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11904
Query Match 2.8%; Score 6; DB 4; Length 155;
Best Local Similarity 100.0%; Pred. No. 2.7e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 GGLSIL 175
Db 47 GGLSIL 52
RESULT 65
US-09-107-532A-4926
; Sequence 4926, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; CORRESPONDENCE ADDRESS:
; ADDRESS: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; OPERATING SYSTEM: <unknown>
; SOFTWARE: ASCII
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4926:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 159 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...159
; SEQUENCE DESCRIPTION: SEQ ID NO: 4926:
US-09-107-532A-4926
Query Match 2.8%; Score 6; DB 4; Length 159;
Best Local Similarity 100.0%; Pred. No. 2.8e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 121 KLGDLH 126
Db 18 KLGDLH 23
RESULT 66
US-09-621-976-4550
; Sequence 4550, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4550
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 150
; OTHER INFORMATION: Xaa = Gly,Arg,Trp
US-09-621-976-4550
Query Match 2.8%; Score 6; DB 4; Length 163;
Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 TKVLAP 204
Db 158 TKVLAP 163
RESULT 67
US-09-489-039A-11750
; Sequence 11750, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11750
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11750
Query Match 2.8%; Score 6; DB 4; Length 164;
Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0; Gaps 0;

```
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVALSL 144
DB 115 VVALSL 120

RESULT 68
US-09-198-452A-1213
; Sequence 1213, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1213
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 1...168
; OTHER INFORMATION: Xaa=unknown or other
US-09-198-452A-1213

Query Match 2.8%; Score 6; DB 4; Length 168;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 SILDPI 178
DB 56 SILDPI 51

RESULT 69
US-08-460-739-3
; Sequence 3, Application US/08460739
; Patent No. 5698416
; GENERAL INFORMATION:
; APPLICANT: Wolf, Marcia K.
; APPLICANT: Cassels, Frederick J.
; APPLICANT: Bell, Brian A.
; TITLE OF INVENTION: Improved Methods for Production of
; TITLE OF INVENTION: Antigens Under Control of Temperature-Regulated Promoters
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Glenna Hendricks
; STREET: 9469 A Main Street
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,739
; FILING DATE: 20-APR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna M.
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: wolf2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 425-4250
; TELEFAX: (703) 425-2767
```

```
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: CPA1 protein
; US-08-460-739-3

Query Match 2.8%; Score 6; DB 1; Length 170;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 FVAVSA 17
DB 16 FVAVSA 21

RESULT 70
US-08-483-101-11
; Sequence 11, Application US/08483101
; Patent No. 5932715
; GENERAL INFORMATION:
; APPLICANT: Scott, June R.
; APPLICANT: Froehlich, Barbara
; APPLICANT: Caron, Judy
; TITLE OF INVENTION: CS2 Proteins and Coding Sequences
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,101
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Terber, Donna M.
; REGISTRATION NUMBER: 33878
; REFERENCE/DOCKET NUMBER: 6-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-483-101-11

Query Match 2.8%; Score 6; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 FVAVSA 17
DB 16 FVAVSA 21
```

```
RESULT 71
US-09-446-959-10
; Sequence 10, Application US/09446959
; Patent No. 6426070
; GENERAL INFORMATION:
; APPLICANT: ROSENBERG, HELENE F.
; APPLICANT: DOMACHOWSKIE, JOSEPH B.
; TITLE OF INVENTION: METHODS FOR INACTIVATING ENVELOPED RNA VIRUS PARTICLES
; TITLE OF INVENTION: AND COMPOSITIONS FOR USE THEREWITH
; FILE REFERENCE: 11613-8USWO
; CURRENT APPLICATION NUMBER: US/09/446,959
; CURRENT FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: PCT/US98/13852
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/052,986
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-446-959-10

Query Match          2.8%; Score 6; DB 4; Length 178;
Best Local Similarity 100.0%; Pred. No. 3.1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY      162 FANVWN 167
      |||||
DB      75 FANVWN 80

RESULT 72
US-09-107-532A-3794
; Sequence 3794, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3794:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 179 amino acids
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...179
; SEQUENCE DESCRIPTION: SEQ ID NO: 3794:
US-09-107-532A-3794

Query Match          2.8%; Score 6; DB 4; Length 179;
Best Local Similarity 100.0%; Pred. No. 3.1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY      169 IGGLSI 174
      |||||
DB      54 IGGLSI 59

RESULT 73
US-09-543-681A-7669
; Sequence 7669, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7669
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7669

Query Match          2.8%; Score 6; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 3.1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY      139 VVALSL 144
      |||||
DB      131 VVALSL 136

RESULT 74
US-09-543-681A-5384
; Sequence 5384, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5384
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5384

Query Match          2.8%; Score 6; DB 4; Length 183;
Best Local Similarity 100.0%; Pred. No. 3.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;
```

```
QY 144 LEISDE 149
DB 105 LEISDE 110

RESULT 75
US-09-230-637-36
; Sequence 36, Application US/09230637
; Patent No. 6264958
; GENERAL INFORMATION:
; APPLICANT: Hayward, Gary
; APPLICANT: Nicholas, John
; APPLICANT: Hardwick, J. Marie
; APPLICANT: Reitz, Marvin
; TITLE OF INVENTION: No. 6264958el Genes of Kaposi's Sarcoma
; FILE REFERENCE: 1107.78372
; CURRENT APPLICATION NUMBER: US/09/230.637
; CURRENT FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: 60/022,591
; PRIOR FILING DATE: 1996-07-25
; PRIOR APPLICATION NUMBER: PCT US 97/12931
; PRIOR FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-230-637-36

Query Match 2.8%; Score 6; DB 3; Length 186;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GVLSDV 185
DB 164 GVLSDV 169

RESULT 76
US-09-134-000C-5358
; Sequence 5358, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134.000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5358
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5358

Query Match 2.8%; Score 6; DB 4; Length 186;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 VALSLE 145
DB 130 VALSLE 135

RESULT 77
US-09-142-530C-7
; Sequence 7, Application US/09142530C
; Patent No. 6642043
; GENERAL INFORMATION:
; APPLICANT: Bertino, Joseph R
; APPLICANT: Erican-Abali, Emine
; APPLICANT: Banerjee, Debabrata
; APPLICANT: Mineishi, Shin
; APPLICANT: Sadelain, Michel
; TITLE OF INVENTION: Double Mutants of Dihydrofolate Reductase and Methods of Using
; FILE REFERENCE: MSK. P-007
; CURRENT APPLICATION NUMBER: US/09/142,530C
; CURRENT FILING DATE: 1999-01-20
; PRIOR APPLICATION NUMBER: PCT/US97/03873
; PRIOR FILING DATE: 1997-03-12
; PRIOR APPLICATION NUMBER: US 60/013,270
; PRIOR FILING DATE: 1996-03-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 186
; TYPE: PRT
; ORGANISM: human
US-09-142-530C-7

Query Match 2.8%; Score 6; DB 4; Length 186;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GVLSDV 185
DB 164 GVLSDV 169

RESULT 78
US-09-198-452A-444
; Sequence 444, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griflais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 444
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-444

Query Match 2.8%; Score 6; DB 4; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 FKGELE 72
DB 92 FKGELE 97

RESULT 79
US-09-134-001C-4319
; Sequence 4319, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
```


; APPLICANT: Ho, Yap
; APPLICANT: Tong, Seow
; TITLE OF INVENTION: METHOD FOR ISOLATION OF BIOSYNTHESIS GENES FOR
; TITLE OF INVENTION: BIOACTIVE MOLECULES
; FILE REFERENCE: 9993-006
; CURRENT APPLICATION NUMBER: US/08/861,774E
; CURRENT FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Parmelia sulcata
US-08-861-774E-44

Query Match 2.8%; Score 6; DB 3; Length 214;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 DDAAIA 39
Db 157 DDAAIA 162

RESULT 84

US-09-143-127-7
; Sequence 7, Application US/09143127
; Patent No. 6605287

; GENERAL INFORMATION:
; APPLICANT: Kousoulas, K.
; APPLICANT: Chouljenko, V.
; APPLICANT: Baghian, A.
; APPLICANT: Tully, Jr., T.
; TITLE OF INVENTION: Vaccines for Chlamydia psittaci
; TITLE OF INVENTION: Infections
; FILE REFERENCE: 21099.0056
; CURRENT APPLICATION NUMBER: US/09/143,127
; CURRENT FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: 60/057,147
; EARLIER FILING DATE: 1997-08-28
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Chlamydia psittaci (cockatiel)
US-09-143-127-7

Query Match 2.8%; Score 6; DB 4; Length 215;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 VLSDVL 186
Db 158 VLSDVL 163

RESULT 85

US-09-143-127-1
; Sequence 1, Application US/09143127
; Patent No. 6605287

; GENERAL INFORMATION:
; APPLICANT: Kousoulas, K.
; APPLICANT: Chouljenko, V.
; APPLICANT: Baghian, A.
; APPLICANT: Tully, Jr., T.
; TITLE OF INVENTION: Vaccines for Chlamydia psittaci
; TITLE OF INVENTION: Infections
; FILE REFERENCE: 21099.0056
; CURRENT APPLICATION NUMBER: US/09/143,127
; CURRENT FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: 60/057,147
; EARLIER FILING DATE: 1997-08-28

; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Chlamydia psittaci (cockatiel)
US-09-143-127-1

Query Match 2.8%; Score 6; DB 4; Length 222;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 VLSDVL 186
Db 165 VLSDVL 170

RESULT 86

US-08-028-463-4
; Sequence 4, Application US/08028463
; Patent No. 5731176

; GENERAL INFORMATION:
; APPLICANT: BEPPU, TERUHIKO
; APPLICANT: YAMADA, HIDEAKI
; APPLICANT: NAGASAWA, TORU
; APPLICANT: HORINOCHI, SUCHAYU
; APPLICANT: NISHIYAMA, MAKATO
; TITLE OF INVENTION: DNA FRAGMENT ENCODING A POLYPEPTIDE
; TITLE OF INVENTION: HAVING NITRILE HYDRATASE ACTIVITY, A TRANSFORMANT CONTAINING THE
; TITLE OF INVENTION: GENE AND A PROCESS FOR THE PRODUCTION OF AMIDES USING THE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: N.Y.
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/028,463
; FILING DATE: 09-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/694,747
; FILING DATE: 02-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7005-023-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-9741
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 226 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Rhodococcus rhodochrous
; STRAIN: J-1 (FERM BP-1478)
US-08-028-463-4

Query Match 2.8%; Score 6; DB 1; Length 226;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 NVNWHI 169
Db 211 NVNWHI 216

RESULT 87
US-08-461-836-4
; Sequence 4, Application US/08461.836
; Patent No. 5753472
; GENERAL INFORMATION:
; APPLICANT: BEPPU, TERUHIKO
; APPLICANT: YAMADA, HIDEAKI
; APPLICANT: NAGASAWA, TORU
; APPLICANT: HORINOUCHI, SUCHAYU
; APPLICANT: NISHIYAMA, MAKATO
; TITLE OF INVENTION: DNA FRAGMENT ENCODING A POLYPEPTIDE
; TITLE OF INVENTION: HAVING NITRILE HYDRATASE ACTIVITY, A TRANSFORMANT CONTAINING T
; TITLE OF INVENTION: GENE AND A PROCESS FOR THE PRODUCTION OF AMIDES USING THE
; TITLE OF INVENTION: TRANSFORMANT
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: N.Y.
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461.836
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/028,463
; FILING DATE: 09-MAR-1993
; APPLICATION NUMBER: US 07/694,747
; FILING DATE: 02-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 19,872
; REFERENCE/DOCKET NUMBER: 7005-023-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-9741
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 226 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Rhodococcus rhodochrous
; STRAIN: J-1 (FERM BP-1478)
; US-08-461-836-4

Query Match 2.8%; Score 6; DB 1; Length 226;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 NVNWHI 169
Db 211 NVNWHI 216

RESULT 88
US-09-489-039A-10204
; Sequence 10204, Application US/09489039A

; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709-2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10204
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-10204

Query Match 2.8%; Score 6; DB 4; Length 230;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 VAFVAV 15
Db 119 VAFVAV 124

RESULT 89
US-09-495-406-35
; Sequence 35, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-00011005
; CURRENT APPLICATION NUMBER: US/09/495,406
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; FEATURE:
; OTHER INFORMATION: putative ORF from GenBank #U32720
; US-09-495-406-35

Query Match 2.8%; Score 6; DB 4; Length 231;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 SWEYDL 118
Db 128 SWEYDL 133

RESULT 90
US-08-557-309B-40
; Sequence 40, Application US/08557309B
; Patent No. 5916572
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; PREVENTION OF T

STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/557.309B
FILING DATE: 14-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.422C1
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 233 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-557-309B-40

Query Match 2.8%; Score 6; DB 2; Length 233;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 FEVRQF 162
DB 46 FEVRQF 51

RESULT 91
US-08-834-306-40
Sequence 40, Application US/08834306
Patent No. 6054135
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond L.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/834,306
FILING DATE: 15-APR-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.422C1
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 40:

SEQUENCE CHARACTERISTICS:
LENGTH: 233 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-834-306-40

Query Match 2.8%; Score 6; DB 3; Length 233;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 FEVRQF 162
DB 46 FEVRQF 51

RESULT 92
US-08-993-674A-40
Sequence 40, Application US/08993674A
Patent No. 6228372
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond L.
APPLICANT: Smith, John M.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,674A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.422C2
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 233 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-993-674A-40

Query Match 2.8%; Score 6; DB 3; Length 233;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 FEVRQF 162
DB 46 FEVRQF 51

RESULT 93
US-09-256-976-40
Sequence 40, Application US/09256976
Patent No. 6419933

GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Smith, John M.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION
; FILE REFERENCE: 210121.422C3
; CURRENT APPLICATION NUMBER: US/09/256,976
; CURRENT FILING DATE: 1999-02-24
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Trypanosoma cruzi
US-09-256-976-40

Query Match 2.8%; Score 6; DB 4; Length 233;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 FEVRQF 162
DB 46 FEVRQF 51

RESULT 94
US-09-374-827-8
; Sequence 8, Application US/09374827
; Patent No. 6603058
; GENERAL INFORMATION:
; APPLICANT: Brennan, Miles B.
; APPLICANT: Hochgeschwender, Ute
; TITLE OF INVENTION: "NON-HUMAN ANIMAL MODEL FOR OBESITY AND USES THEREOF"
; FILE REFERENCE: 3718-5
; CURRENT APPLICATION NUMBER: US/09/374,827
; CURRENT FILING DATE: 1999-08-12
; EARLIER APPLICATION NUMBER: 60/111,581
; EARLIER FILING DATE: 1998-12-09
; EARLIER APPLICATION NUMBER: 60/146,306
; EARLIER FILING DATE: 1998-07-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-374-827-8

Query Match 2.8%; Score 6; DB 4; Length 235;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 FXRELE 211
DB 162 FXRELE 167

RESULT 95
US-09-252-991A-28804
; Sequence 28804, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28714
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28714

Query Match 2.8%; Score 6; DB 4; Length 240;

GENERAL INFORMATION:
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28804
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28804

Query Match 2.8%; Score 6; DB 4; Length 236;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 DAIAAI 40
DB 74 DAIAAI 79

RESULT 96
US-09-252-991A-29211
; Sequence 29211, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29211
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29211

Query Match 2.8%; Score 6; DB 4; Length 237;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 QGDANV 92
DB 68 QGDANV 73

RESULT 97
US-09-252-991A-28714
; Sequence 28714, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28714
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28714

Query Match 2.8%; Score 6; DB 4; Length 240;

Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 GELAMR 74
|||||
Db 51 GELAMR 56

RESULT 98
US-09-522-714-20
; Sequence 20, Application US/09522714
; Patent No. 6563020
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Valpani, Nasser
; TITLE OF INVENTION: Maize Chitinases and Their Use in
; TITLE OF INVENTION: Enhancing Disease Resistance in Crop Plants
; FILE REFERENCE: 1100
; CURRENT APPLICATION NUMBER: US/09/522,714
; CURRENT FILING DATE: 2000-03-10
; EARLIER APPLICATION NUMBER: 60/125,915
; EARLIER FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Zea mays
US-09-522-714-20

Query Match 2.8%; Score 6; DB 4; Length 252;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 EEINKA 32
|||||
Db 111 EEINKA 116

Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 GELAMR 74
|||||
Db 51 GELAMR 56

RESULT 98
US-09-522-714-20
; Sequence 20, Application US/09522714
; Patent No. 6563020
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Valpani, Nasser
; TITLE OF INVENTION: Maize Chitinases and Their Use in
; TITLE OF INVENTION: Enhancing Disease Resistance in Crop Plants
; FILE REFERENCE: 1100
; CURRENT APPLICATION NUMBER: US/09/522,714
; CURRENT FILING DATE: 2000-03-10
; EARLIER APPLICATION NUMBER: 60/125,915
; EARLIER FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Zea mays
US-09-522-714-20

Query Match 2.8%; Score 6; DB 4; Length 252;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 EEINKA 32
|||||
Db 111 EEINKA 116

Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 IAAVAF 12
|||||
Db 77 IAAVAF 82

Search completed: August 6, 2004, 16:10:32
Job time : 20 secs

Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 EEINKA 32
|||||
Db 111 EEINKA 116

RESULT 99
US-09-489-039A-13690
; Sequence 13690, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13690
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13690

Query Match 2.8%; Score 6; DB 4; Length 252;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LIIAAV 10
|||||
Db 30 LIIAAV 35

Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 IAAVAF 12
|||||
Db 77 IAAVAF 82

Search completed: August 6, 2004, 16:10:32
Job time : 20 secs

Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 IAAVAF 12
|||||
Db 77 IAAVAF 82

Search completed: August 6, 2004, 16:10:32
Job time : 20 secs

Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LIIAAV 10
|||||
Db 30 LIIAAV 35

RESULT 100
US-09-543-681A-7108
; Sequence 7108, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2004, 16:09:43 ; Search time 46 Seconds
(without alignments)
1452.488 Million cell updates/sec

Title: US-10-024-955-7

Perfect score: 213

Sequence: 1 MMKFLIIAAVFAVVSADPI.....VKEMTKVLAPFAKRELEKN 213

Scoring table: OLIGO

Gapop 60.0 , Gapext 50.0

Searched: 1291235 seqs, 313682936 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1291235

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	213	100.0	213	10	US-09-847-208-77
2	213	100.0	213	13	US-10-024-955-7
3	25	11.7	215	10	US-09-847-208-85
4	25	11.7	215	13	US-10-024-955-2
5	8	3.8	484	12	US-10-282-122A-51279
6	8	3.8	611	12	US-10-670-695-32
7	7	3.3	41	15	US-10-283-940-19
8	7	3.3	62	12	US-10-424-599-151546
9	7	3.3	87	9	US-09-815-242-13280
10	7	3.3	87	12	US-10-282-122A-72614
11	7	3.3	87	12	US-10-282-122A-73989
12	7	3.3	87	12	US-10-282-122A-74627
13	7	3.3	105	12	US-10-424-599-168513
14	7	3.3	124	12	US-10-424-599-161131
15	7	3.3	146	12	US-10-424-599-175474

16	7	3.3	166	16	US-10-437-963-198443	Sequence 198443, App
17	7	3.3	169	12	US-10-425-114-48247	Sequence 48247, A
18	7	3.3	188	12	US-10-424-599-233048	Sequence 233048, App
19	7	3.3	204	14	US-10-316-253-305	Sequence 305, App
20	7	3.3	219	12	US-10-424-599-175476	Sequence 175476, App
21	7	3.3	245	16	US-10-437-963-165041	Sequence 165041, App
22	7	3.3	253	12	US-10-424-599-179367	Sequence 179367, App
23	7	3.3	266	12	US-10-424-599-164461	Sequence 164461, App
24	7	3.3	267	15	US-10-104-047-2248	Sequence 2248, App
25	7	3.3	277	12	US-10-425-114-57308	Sequence 57308, A
26	7	3.3	295	14	US-10-205-194-95	Sequence 95, Appl
27	7	3.3	297	16	US-10-437-963-176814	Sequence 176814, App
28	7	3.3	299	12	US-10-424-599-163477	Sequence 163477, App
29	7	3.3	362	12	US-10-424-599-276435	Sequence 276435, App
30	7	3.3	368	12	US-10-282-122A-63339	Sequence 63339, A
31	7	3.3	372	12	US-10-282-122A-68363	Sequence 68363, A
32	7	3.3	406	14	US-10-156-761-14261	Sequence 14261, A
33	7	3.3	422	16	US-10-437-963-102813	Sequence 102813, App
34	7	3.3	430	12	US-10-282-122A-51513	Sequence 51513, A
35	7	3.3	453	12	US-10-282-122A-52319	Sequence 52319, A
36	7	3.3	468	12	US-10-282-122A-52924	Sequence 52924, A
37	7	3.3	486	12	US-10-425-114-39247	Sequence 39247, A
38	7	3.3	486	12	US-10-425-114-50309	Sequence 50309, A
39	7	3.3	487	12	US-10-425-114-68137	Sequence 68137, A
40	7	3.3	498	14	US-10-193-896-5	Sequence 5, Appl
41	7	3.3	518	12	US-10-425-114-68294	Sequence 68294, A
42	7	3.3	538	15	US-10-369-493-18179	Sequence 18179, A
43	7	3.3	542	16	US-10-437-963-165038	Sequence 165038, App
44	7	3.3	551	12	US-10-424-599-151822	Sequence 151822, App
45	7	3.3	581	15	US-10-094-466-40	Sequence 40, Appl
46	7	3.3	595	15	US-10-369-493-17076	Sequence 17076, A
47	7	3.3	625	12	US-10-282-122A-48336	Sequence 48336, A
48	7	3.3	718	14	US-10-017-161-784	Sequence 784, App
49	7	3.3	719	16	US-10-437-963-107435	Sequence 107435, App
50	7	3.3	763	16	US-10-437-963-173899	Sequence 173899, App
51	7	3.3	777	16	US-10-437-963-173900	Sequence 173900, App
52	7	3.3	913	15	US-10-369-493-7255	Sequence 7255, App
53	7	3.3	924	15	US-10-369-493-4496	Sequence 4496, App
54	7	3.3	1001	12	US-10-282-122A-49603	Sequence 49603, A
55	6	2.8	10	12	US-10-014-340-317	Sequence 317, Appl
56	6	2.8	12	10	US-09-862-955C-20	Sequence 20, Appl
57	6	2.8	15	9	US-09-073-009-73	Sequence 73, Appl
58	6	2.8	15	9	US-09-073-009-75	Sequence 75, Appl
59	6	2.8	15	9	US-09-793-306-73	Sequence 73, Appl
60	6	2.8	15	9	US-09-793-306-75	Sequence 75, Appl
61	6	2.8	19	14	US-10-215-435-3	Sequence 3, Appl
62	6	2.8	19	16	US-10-822-072-3	Sequence 3, Appl
63	6	2.8	21	9	US-09-864-761-36329	Sequence 36329, A
64	6	2.8	28	14	US-10-360-053-23	Sequence 23, Appl
65	6	2.8	32	12	US-10-424-599-254894	Sequence 254894, App
66	6	2.8	33	9	US-09-864-761-41249	Sequence 41249, A
67	6	2.8	37	12	US-10-424-599-172112	Sequence 172112, App
68	6	2.8	42	12	US-10-424-599-245873	Sequence 245873, App
69	6	2.8	43	9	US-09-864-761-40280	Sequence 40280, A
70	6	2.8	43	12	US-10-424-599-219711	Sequence 219711, App
71	6	2.8	45	12	US-10-424-599-257467	Sequence 257467, App
72	6	2.8	49	12	US-10-424-599-188617	Sequence 188617, App
73	6	2.8	51	12	US-10-424-599-182164	Sequence 182164, App
74	6	2.8	53	16	US-10-437-963-142935	Sequence 142935, App
75	6	2.8	54	12	US-10-424-599-148141	Sequence 148141, App
76	6	2.8	55	12	US-10-424-599-266220	Sequence 266220, App
77	6	2.8	56	12	US-10-424-599-236065	Sequence 236065, App
78	6	2.8	56	15	US-10-264-049-3669	Sequence 3669, App
79	6	2.8	57	14	US-10-211-962-124	Sequence 124, App
80	6	2.8	57	16	US-10-437-963-188318	Sequence 188318, App
81	6	2.8	58	9	US-09-864-761-35621	Sequence 35621, A
82	6	2.8	59	12	US-10-424-599-143933	Sequence 143933, App
83	6	2.8	59	12	US-10-424-599-252348	Sequence 252348, App
84	6	2.8	60	12	US-10-424-599-245904	Sequence 245904, App
85	6	2.8	61	16	US-10-437-963-197055	Sequence 197055, App
86	6	2.8	64	12	US-10-424-599-285188	Sequence 285188, App
87	6	2.8	65	12	US-10-424-599-244365	Sequence 244365, App
88	6	2.8	65	16	US-10-437-963-108901	Sequence 108901, App

89 Sequence 119979,
90 Sequence 143189,
91 Sequence 228585,
92 Sequence 163372,
93 Sequence 184690,
94 Sequence 179217,
95 Sequence 6222, Ap
96 Sequence 181866,
97 Sequence 227078,
98 Sequence 282630,
99 Sequence 151004,
100 Sequence 40082, A

6 2.8 65 16 US-10-437-963-113979
6 2.8 65 16 US-10-437-963-143189
6 2.8 66 12 US-10-424-599-228585
6 2.8 66 16 US-10-437-963-163372
6 2.8 66 16 US-10-437-963-184690
6 2.8 67 12 US-10-424-599-179217
6 2.8 69 9 US-09-738-626-6222
6 2.8 69 12 US-10-424-599-181866
6 2.8 72 12 US-10-424-599-227078
6 2.8 72 12 US-10-424-599-282630
6 2.8 74 12 US-10-424-599-151004
6 2.8 76 12 US-10-425-114-40082

ALIGNMENTS

RESULT 1

US-09-847-208-77
; Sequence 77, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daoheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES
; FILE REFERENCE: UC67.002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae (House-dust mite)
US-09-847-208-77

Query Match 100.0%; Score 213; DB 10; Length 213;
Best Local Similarity 100.0%; Pred. No. 2.1e-201;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKKFLIIAAVAFVAVSADPHYDKITTEINKAIDDAIAAEQSETIDPMKVPDHADKFER 60
DB 1 MKKFLIIAAVAFVAVSADPHYDKITTEINKAIDDAIAAEQSETIDPMKVPDHADKFER 60
QY 61 HVGIVDFKGLAMRNIEARGLKQMKRQGDANVKGEIGIVKAHLIGVHDDIVSMEYDLAY 120
DB 61 HVGIVDFKGLAMRNIEARGLKQMKRQGDANVKGEIGIVKAHLIGVHDDIVSMEYDLAY 120
QY 121 KLGLHPHTHVISDIQDFVVALSLEISDEGNITMTSEVRFQFANVNVNHIGLSILDPFG 180
DB 121 KLGLHPHTHVISDIQDFVVALSLEISDEGNITMTSEVRFQFANVNVNHIGLSILDPFG 180
QY 181 VLSDVLTAFQDVTVRKEMTKVLAFAFKRELEKN 213
DB 181 VLSDVLTAFQDVTVRKEMTKVLAFAFKRELEKN 213

RESULT 2

US-10-024-955-7
; Sequence 7, Application US/10024955
; Publication No. US20020168373A1
; GENERAL INFORMATION:
; APPLICANT: Wayne R. Thomas and Kaw-Yan Chua
; TITLE OF INVENTION: Allergenic Proteins and Peptides From
; TITLE OF INVENTION: House Dust Mite and Uses Therefor
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts

COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/10/024,955
APPLICATION NUMBER: US/10/024,955
FILING DATE: 19-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/553,336A
FILING DATE: 10-JUN-1996
APPLICATION NUMBER: US 08/081,540
FILING DATE: 22-JUNE-1993
ATTORNEY/AGENT INFORMATION:
NAME: Jane E. Remillard
REGISTRATION NUMBER: 38,872
REFERENCE/DOCKET NUMBER: IMI-032CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 213 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Query Match 100.0%; Score 213; DB 13; Length 213;
Best Local Similarity 100.0%; Pred. No. 2.1e-201;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKKFLIIAAVAFVAVSADPHYDKITTEINKAIDDAIAAEQSETIDPMKVPDHADKFER 60
DB 1 MKKFLIIAAVAFVAVSADPHYDKITTEINKAIDDAIAAEQSETIDPMKVPDHADKFER 60
QY 61 HVGIVDFKGLAMRNIEARGLKQMKRQGDANVKGEIGIVKAHLIGVHDDIVSMEYDLAY 120
DB 61 HVGIVDFKGLAMRNIEARGLKQMKRQGDANVKGEIGIVKAHLIGVHDDIVSMEYDLAY 120
QY 121 KLGLHPHTHVISDIQDFVVALSLEISDEGNITMTSEVRFQFANVNVNHIGLSILDPFG 180
DB 121 KLGLHPHTHVISDIQDFVVALSLEISDEGNITMTSEVRFQFANVNVNHIGLSILDPFG 180
QY 181 VLSDVLTAFQDVTVRKEMTKVLAFAFKRELEKN 213
DB 181 VLSDVLTAFQDVTVRKEMTKVLAFAFKRELEKN 213

RESULT 3

US-09-847-208-85
; Sequence 85, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daoheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES
; FILE REFERENCE: UC67.002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 85
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus (House-dust mite)
US-09-847-208-85

Query Match 11.7%; Score 25; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 3.4e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 TSFEVQFANVNHIGLSILDPIF 179
Db 155 TSFEVQFANVNHIGLSILDPIF 179

RESULT 4

US-10-024-955-2
; Sequence 2, Application US/10024955
; Publication No. US20020168373A1
; GENERAL INFORMATION:
; APPLICANT: Wayne R. Thomas and Kaw-Yan Chua
; TITLE OF INVENTION: Allergenic Proteins and Peptides From
; House Dust Mite and Uses Therefor
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/024,955
; FILING DATE: 19-Dec-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/553,336A
; FILING DATE: 10-JUN-1996
; APPLICATION NUMBER: US 08/081,540
; FILING DATE: 22-JUNE-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane E. Remillard
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: IMI-032CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-024-955-2

Query Match 11.7%; Score 25; DB 13; Length 215;
Best Local Similarity 100.0%; Pred. No. 3.4e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 TSFEVQFANVNHIGLSILDPIF 179
Db 155 TSFEVQFANVNHIGLSILDPIF 179

RESULT 5

US-10-282-122A-51279
; Sequence 51279, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51279
; TYPE: PRT
; LENGTH: 484
; ORGANISM: Bordetella pertussis
US-10-282-122A-51279

Query Match 3.8%; Score 8; DB 12; Length 484;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 IEARGLKQ 83
Db 415 IEARGLKQ 422

RESULT 6

US-10-670-695-32
; Sequence 32, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Feline leukemia virus

US-10-670-695-32

Query Match 3.8%; Score 8; DB 12; Length 611;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 KEMTKVLA 203
|||||
Db 131 KEMTKVLA 138

RESULT 7

US-10-283-940-19
; Sequence 19, Application US/10283940
; Publication No. US20030220394A1
; GENERAL INFORMATION:
; APPLICANT: Danisco A/S
; APPLICANT: Morgan, Andrew J
; APPLICANT: Yu, Shukun
; APPLICANT: Wei, Gang, Inge
; APPLICANT: Pedersen, Hans C
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 674509-2041.1
; CURRENT APPLICATION NUMBER: US/10/283,940
; CURRENT FILING DATE: 2002-10-30
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: PCT/GB02/04916
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: GB 0126164.3
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US 60/343,485
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 19
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Phanerochaete chrysosporium
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (19)-(19)
; OTHER INFORMATION: Xaa is an unknown amino acid residue
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (19)-(19)
; OTHER INFORMATION: Xaa is Ser in the database sequence
US-10-283-940-19

Query Match 3.3%; Score 7; DB 15; Length 41;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 SDIQDFV 139
|||||
Db 1 SDIQDFV 7

RESULT 8

US-10-424-599-151546
; Sequence 151546, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 151546
; LENGTH: 62

; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MPT3847_107869C.1.pep
US-10-424-599-151546

Query Match 3.3%; Score 7; DB 12; Length 62;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 FLIIAAV 10
|||||
Db 2 FLIIAAV 8

RESULT 9

US-09-815-242-13280
; Sequence 13280, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13280
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13280

Query Match 3.3%; Score 7; DB 9; Length 87;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 DDAAIAAI 40
|||||
Db 71 DDAAIAAI 77

RESULT 10

US-10-282-122A-72614
; Sequence 72614, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malore, Cheryl
; APPLICANT: Haselbeck, Robert


```
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72614
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-282-122A-72614

Query Match          3.3%; Score 7; DB 12; Length 87;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      34 DDATAAI 40
DB      71 DDATAAI 77

RESULT 11
US-10-282-122A-73989
; Sequence 73989, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72614
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-282-122A-72614
```

```
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73989
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73989

Query Match          3.3%; Score 7; DB 12; Length 87;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      34 DDATAAI 40
DB      71 DDATAAI 77

RESULT 12
US-10-282-122A-74627
; Sequence 74627, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73989
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73989
```

```
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74627
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-282-122A-74627

Query Match      3.3%; Score 7; DB 12; Length 87;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 DAIAAIE 41
Db 72 DAIAAIE 78

RESULT 13
US-10-424-599-168513
; Sequence 168513, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 168513
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_123182C.1.pep
US-10-424-599-168513

Query Match      3.3%; Score 7; DB 12; Length 105;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 VHDDIVS 113
Db 90 VHDDIVS 96

RESULT 14
US-10-424-599-161131
; Sequence 161131, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 161131
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Glycine max
US-10-424-599-161131

; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_116519C.1.pep
US-10-424-599-161131

Query Match      3.3%; Score 7; DB 12; Length 124;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 AIAAIEQ 42
Db 64 AIAAIEQ 70

RESULT 15
US-10-424-599-175474
; Sequence 175474, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 175474
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_129471C.1.pep
US-10-424-599-175474

Query Match      3.3%; Score 7; DB 12; Length 146;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 LSLEISD 148
Db 43 LSLEISD 49

RESULT 16
US-10-437-963-198443
; Sequence 198443, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 198443
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_94101C.1.pep
US-10-437-963-198443

Query Match      3.3%; Score 7; DB 16; Length 166;
```

Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 185 VLTAFQ 191
Db 79 VLTAFQ 85

RESULT 17

US-10-425-114-48247
; Sequence 48247, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 48247
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3600-046-A9_FLI.pep
US-10-425-114-48247

Query Match 3.3%; Score 7; DB 12; Length 169;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 185 VLTAFQ 191
Db 84 VLTAFQ 90

RESULT 18

US-10-424-599-233048
; Sequence 233048, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 233048
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_52467C.1.pep
US-10-424-599-233048

Query Match 3.3%; Score 7; DB 12; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 185 VLTAFQ 191
Db 91 VLTAFQ 97

RESULT 19

US-10-316-253-306
; Sequence 306, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 306
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-316-253-306

Query Match 3.3%; Score 7; DB 14; Length 204;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSADP 19
Db 59 VAVSADP 65

RESULT 20

US-10-424-599-175476
; Sequence 175476, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 175476
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_129473C.1.pep
US-10-424-599-175476

Query Match 3.3%; Score 7; DB 12; Length 219;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 LSLEISD 148
Db 116 LSLEISD 122

RESULT 21

US-10-437-963-165041
; Sequence 165041, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 165041
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_63885C.1.pep
US-10-437-963-165041

Query Match 3.3%; Score 7; DB 16; Length 245;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 FLLIAAV 10
DB 11 FLLIAAV 17

RESULT 22
US-10-424-599-179367
; Sequence 179367, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 179367
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_132985C.1.pep
US-10-424-599-179367

Query Match 3.3%; Score 7; DB 12; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 SILDPPIF 179
DB 110 SILDPPIF 116

RESULT 23
US-10-424-599-164461
; Sequence 164461, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 164461
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_119526C.1.pep
US-10-424-599-164461

Query Match 3.3%; Score 7; DB 12; Length 266;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 EQSETID 47
DB 152 EQSETID 158

RESULT 24
US-10-104-047-2248
; Sequence 2248, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2248
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2248

Query Match 3.3%; Score 7; DB 15; Length 267;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 DDAAIAI 40
DB 45 DDAAIAI 51

RESULT 25
US-10-425-114-57308
; Sequence 57308, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 57308
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17258G10_FLI.pep
US-10-425-114-57308

Query Match 3.3%; Score 7; DB 12; Length 277;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 IGGLSIL 175
DB 139 IGGLSIL 145

RESULT 26
US-10-205-194-95
; Sequence 95, Application US/10205194
; Publication No. US20030134301A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pirneock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018201
; CURRENT APPLICATION NUMBER: US/10/205,194
; CURRENT FILING DATE: 5200-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 95
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Beta chimaerin
US-10-205-194-95

Query Match 3.3%; Score 7; DB 14; Length 295;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 IEARGLK 82
DB 128 IEARGLK 134

RESULT 27
US-10-437-963-176814
; Sequence 176814, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 176814
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_74527C.1.pep
US-10-437-963-176814

Query Match 3.3%; Score 7; DB 16; Length 297;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LLIAAVA 11
DB 11 LLIAAVA 17

RESULT 28
US-10-424-599-163477
; Sequence 163477, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 163477
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(299)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_118639C.1.pep
US-10-424-599-163477

Query Match 3.3%; Score 7; DB 12; Length 299;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 LSILDPI 178
DB 213 LSILDPI 219

RESULT 29
US-10-424-599-276435
; Sequence 276435, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 276435
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(362)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_9163C.1.pep
US-10-424-599-276435

Query Match 3.3%; Score 7; DB 12; Length 362;
Best Local Similarity 100.0%; Pred. No. 3e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 SLEISDE 149
 Db 166 SLEISDE 172

RESULT 30

US-10-282-122A-63339
 ; Sequence 63339, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 63339

; LENGTH: 368

; TYPE: PRT

; ORGANISM: Moraxella catarrhalis

US-10-282-122A-63339

Query Match 3.3%; Score 7; DB 12; Length 368;

Best Local Similarity 100.0%; Pred.No. 3.1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 PTHVIS 133
 Db 79 PTHVIS 85

RESULT 31

US-10-282-122A-68363
 ; Sequence 68363, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangu
 ; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 68363

; LENGTH: 372

; TYPE: PRT

; ORGANISM: Pseudomonas putida

US-10-282-122A-68363

Query Match 3.3%; Score 7; DB 12; Length 372;

Best Local Similarity 100.0%; Pred.No. 3.1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 TKVLAPA 205
 Db 268 TKVLAPA 274

RESULT 32

US-10-156-761-14261
 ; Sequence 14261, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 14261
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14261

Query Match 3.3%; Score 7; DB 14; Length 406;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 5 LLIAAVA 11
| | | | |
Db 387 LLIAAVA 393

RESULT 33

US-10-437-963-102813
; Sequence 102813, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 102813
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_100301C.1.pap
US-10-437-963-102813

Query Match 3.3%; Score 7; DB 16; Length 422;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 180 GVLSDDL 186
| | | | |
Db 390 GVLSDDL 396

RESULT 34

US-10-282-122A-51513
; Sequence 51513, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51513
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
US-10-282-122A-51513

Query Match 3.3%; Score 7; DB 12; Length 430;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 YDKITEE 28
| | | | |
Db 192 YDKITEE 198

RESULT 35

US-10-282-122A-52319
; Sequence 52319, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931

Query Match 3.3%; Score 7; DB 12; Length 453;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 52319
LENGTH: 453
TYPE: PRT
ORGANISM: Clostridium botulinum
US-10-282-122A-52319

Query Match 3.3%; Score 7; DB 12; Length 453;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 DLAYKLG 123
Db 25 DLAYKLG 31

RESULT 36
US-10-282-122A-52924
; Sequence 52924, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 52924
LENGTH: 468
TYPE: PRT
ORGANISM: Clostridium difficile
US-10-282-122A-52924

Query Match 3.3%; Score 7; DB 12; Length 468;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 117 DLAYKLG 123
Db 41 DLAYKLG 47

RESULT 37
US-10-425-114-39247
; Sequence 39247, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 39247
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3060-084-H1_FLI.pep
US-10-425-114-39247

Query Match 3.3%; Score 7; DB 12; Length 486;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GVLSDVL 186
Db 454 GVLSDVL 460

RESULT 38
US-10-425-114-50309
; Sequence 50309, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 50309
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3060-015-H1_FLI.pep
US-10-425-114-50309

Query Match 3.3%; Score 7; DB 12; Length 486;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GVLSDVL 186

Db 454 GVLSDVL 460
|||||||

RESULT 39

US-10-425-114-68137
; Sequence 68137, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 68137
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17220D07_FLI.pep
US-10-425-114-68137

Query Match 3.3%; Score 7; DB 12; Length 487;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GVLSDVL 186
|||||||
Db 455 GVLSDVL 461

RESULT 40

US-10-193-896-5
; Sequence 5, Application US/10193896
; Publication No. US20030129710A1
; GENERAL INFORMATION:
; APPLICANT: Biotechnologisk Institut
; APPLICANT: Jorgensen, Flemming
; APPLICANT: Hansen, Ole C.
; APPLICANT: Stougaard, Peter
; APPLICANT: Berthelsen, Hans
; APPLICANT: Eriksnaer, Kristian
; APPLICANT: Bottcher, Karen
; APPLICANT: Christensen, Hans Jorgen Singel
; TITLE OF INVENTION: A novel thermostable isomerase and use
; TITLE OF INVENTION: hereof
; FILE REFERENCE: 30077US02
; CURRENT FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: US/10/193,896
; CURRENT FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: 60/305,155
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/905,108
; PRIOR FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Y.pestis
US-10-193-896-5

Query Match 3.3%; Score 7; DB 14; Length 498;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 IGLSIL 175

Db 90 IGLSIL 96
|||||||

RESULT 41

US-10-425-114-68294
; Sequence 68294, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 68294
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73279G08_FLI.pep
US-10-425-114-68294

Query Match 3.3%; Score 7; DB 12; Length 518;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 FLLIAAV 10
|||||||
Db 117 FLLIAAV 123

RESULT 42

US-10-369-493-18179
; Sequence 18179, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18179
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Thermoplasma acidophilum
US-10-369-493-18179

Query Match 3.3%; Score 7; DB 15; Length 538;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 IDDAIAA 39
|||||||
Db 270 IDDAIAA 276

RESULT 43

US-10-437-963-165038

```
; Sequence 165038, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 165038
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_63882C.1.pep
US-10-437-963-165038

Query Match      3.3%; Score 7; DB 16; Length 542;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 FLIIAAV 10
      |||||
DB      11 FLIIAAV 17

RESULT 44
US-10-424-599-151822
; Sequence 151822, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 151822
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_108119C.1.pep
US-10-424-599-151822

Query Match      3.3%; Score 7; DB 12; Length 551;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      31 KAIDDAI 37
      |||||
DB      57 KAIDDAI 63

RESULT 45
US-10-094-466-40
; Sequence 40, Application US/10094466
; Publication No. US20030203363A1
; GENERAL INFORMATION:
; APPLICANT: Spytek et al.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
```

```
; TITLE OF INVENTION: AND METHODS OF USING
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 21402-290D
; CURRENT APPLICATION NUMBER: US/10/094,466
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/288,148
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/338,375
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/275,579
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/335,302
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/275,601
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/276,000
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/277,338
; PRIOR FILING DATE: 2001-03-20
; PRIOR Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patin 2.1
; SEQ ID NO 40
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-466-40

Query Match      3.3%; Score 7; DB 15; Length 581;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      34 DDAAIAAI 40
      |||||
DB      289 DDAAIAAI 295

RESULT 46
US-10-369-493-17076
; Sequence 17076, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17076
; LENGTH: 595
; TYPE: PRT
; ORGANISM: Caulobacter crescentus
US-10-369-493-17076

Query Match      3.3%; Score 7; DB 15; Length 595;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 LLIAAIVA 11
      |||||
```

Db 177 LLIAAVA 183

RESULT 47

US-10-282-1222A-48336

Sequence 48336, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John

APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert

APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 48336

LENGTH: 625

TYPE: PRT

ORGANISM: Bacteroides fragilis

US-10-282-122A-48336

Query Match 3.3%; Score 7; DB 12; Length 625;

Best Local Similarity 100.0%; Pred. No. 5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 PAFKREL 210

Db 525 PAFKREL 531

RESULT 48

US-10-017-161-784

Sequence 784, Application US/10017161

Publication No. US20030143668A1

GENERAL INFORMATION:

APPLICANT: SUWA, MAKIKO

APPLICANT: ASAI, KIYOSHI

APPLICANT: AKIYAMA, YUTAKA

APPLICANT: ABRUTANI, HIROYUKI

TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS

FILE REFERENCE: 084335/0152

US-10-017-161-784

Query Match 3.3%; Score 7; DB 12; Length 625;

Best Local Similarity 100.0%; Pred. No. 5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 PAFKREL 210

Db 525 PAFKREL 531

RESULT 49

US-10-437-963-107435

Sequence 107435, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 107435

LENGTH: 719

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_11788C.1.pep

US-10-437-963-107435

Query Match 3.3%; Score 7; DB 16; Length 719;

Best Local Similarity 100.0%; Pred. No. 5.7e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 KAIDDAI 37

Db 240 KAIDDAI 246

RESULT 50

US-10-437-963-173899

Sequence 173899, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53221)B

US-10-437-963-173899

Query Match 3.3%; Score 7; DB 16; Length 719;

Best Local Similarity 100.0%; Pred. No. 5.7e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 KAIDDAI 37

Db 240 KAIDDAI 246

; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 173899
; LENGTH: 763
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(763)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_71893C.1.pep
US-10-437-963-173899

Query Match 3.3%; Score 7; DB 16; Length 763;
Best Local Similarity 100.0%; Pred. No. 6.e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AVSADPI 20
Db 654 AVSADPI 660

RESULT 51
US-10-437-963-173900
; Sequence 173900, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 173900
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_71894C.1.pep
US-10-437-963-173900

Query Match 3.3%; Score 7; DB 16; Length 777;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AVSADPI 20
Db 668 AVSADPI 674

RESULT 52
US-10-369-493-7255
; Sequence 7255, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 7255
; LENGTH: 913
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(913)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-7255

Query Match 3.3%; Score 7; DB 15; Length 913;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AVSADPI 20
Db 172 AVSADPI 178

RESULT 53
US-10-369-493-4496
; Sequence 4496, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 4496
; LENGTH: 924
; TYPE: PRT
; ORGANISM: Burkholderia fungorum
US-10-369-493-4496

Query Match 3.3%; Score 7; DB 15; Length 924;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AVSADPI 20
Db 168 AVSADPI 174

RESULT 54
US-10-282-122A-49603
; Sequence 49603, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haseibeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.

```

1  APPLICANT: XU, H.
2  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
3  FILE REFERENCE: ELITRA.034A
4  CURRENT APPLICATION NUMBER: US/10/282,122A
5  PRIORITY FILING DATE: 2003-02-20
6  PRIOR APPLICATION NUMBER: 60/191,078
7  PRIOR FILING DATE: 2000-03-21
8  PRIOR APPLICATION NUMBER: 60/206,848
9  PRIOR FILING DATE: 2000-05-23
10 PRIOR APPLICATION NUMBER: 60/207,727
11 PRIOR FILING DATE: 2000-05-26
12 PRIOR APPLICATION NUMBER: 60/230,335
13 PRIOR FILING DATE: 2000-09-06
14 PRIOR APPLICATION NUMBER: 60/230,347
15 PRIOR FILING DATE: 2000-09-09
16 PRIOR APPLICATION NUMBER: 60/242,578
17 PRIOR FILING DATE: 2000-10-23
18 PRIOR APPLICATION NUMBER: 60/253,625
19 PRIOR FILING DATE: 2000-11-27
20 PRIOR APPLICATION NUMBER: 60/257,931
21 PRIOR FILING DATE: 2000-12-22
22 PRIOR APPLICATION NUMBER: 60/267,636
23 PRIOR FILING DATE: 2001-02-09
24 PRIOR APPLICATION NUMBER: 60/269,308
25 PRIOR FILING DATE: 2001-02-16
26 Remaining Prior Application data removed - See File Wrapper or PALM.
27 NUMBER OF SEQ ID NOS: 78614
28 SOFTWARE: PatentIn version 3.1
29 SEQ ID NO 49603
30 LENGTH: 1001
31 TYPE: PRT
32 ORGANISM: Burkholderia fungorum
33 US-10-282-122A-49603

```

Query Match 3.3%; Score 7; DB 12; Length 1001;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	14	AVSADPI	20
D _b	181	AVSADPI	187

```

RESULT 55
US-10-014-340-317
; Sequence 317, Application US/10014340
; Publication NO. US2003006411A1
; GENERAL INFORMATION:
; APPLICANT: Herath, et al
; TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and Uses Therefor, Including
; : Diagnosis and Treatment of Alzheimer's Disease
; FILE REFERENCE: 9195-078
; CURRENT APPLICATION NUMBER: US/10/014,340
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 317
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-014-340-317

```

```
Query Match          2.8%; Score 6; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	102 HLLIGV 107
Db	1 HLLIGV 6

RESULT 56
US-09-962-955C-20

```

: Sequence 20, Application US/09962955C
: Publication No. US20030013648A1
: GENERAL INFORMATION:
: APPLICANT: Gerardo M. Castillo
: APPLICANT: Alan D. Snow
: NUMBER OF SEQUENCES: 37
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Patrick M. Dwyer
: STREET: Proteotech, Inc, 1818 Westlake Avenue N, Suite 114
: CITY: Seattle
: STATE: WA (Washington)
: COUNTRY: United States of America
: ZIP: 98109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette - 3.50 inch, 1.44 Mb storage
: COMPUTER: IBM PC
: OPERATING SYSTEM: Windows 98
: SOFTWARE: Wordperfect 9
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/962,955C
: FILING DATE: 24-September-2001
: CLASSIFICATION:
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 09/938,275
: FILING DATE: 22-August-2001
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Dwyer, Patrick M.
: REGISTRATION NUMBER: 32,411
: REFERENCE/DOCKET NUMBER: PROTEO.P03C1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 343-7074
: TELEFAX: (206) 343-7085
: INFORMATION FOR SEQ ID NO: 20:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 12 AMINO ACIDS
: TYPE: AMINO ACID
: STRANDEDNESS:
: TOPOLOGY: LINEAR
: ORIGINAL SOURCE:
: ORGANISM: MOUSE
: FEATURE:
: OTHER INFORMATION: Also referred to in the specification as
: US-09-962-955C-20

```

Query Match 2.8%; Score 6; DB 10; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 VASAD 18
Db 6 VASAD 11

```

RESULT 57
US-09-073-009-73
; Sequence 73, Application US/09073009
; Patent No. US2001001288A1
; GENERAL INFORMATION:
; APPLICANT: Alderson, Mark
; APPLICANT: Dillon, David C.
; APPLICANT: Skeiky, Yagir A.W.
; APPLICANT: Campos-Neto, Antonio
; TITLE OF INVENTION: COMPOUNDS AND DIAGNOSIS OF
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY
; STREET: 6300 Cumbia Center, 701 Fifth Ave.
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,009
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.441C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
US-09-073-009-73

Query Match 2.8%; Score 6; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 183 SDVLTA 188
Db 8 SDVLTA 13

RESULT 58
US-09-073-009-75
Sequence 75, Application US/09073009
Patent No. US20010012888A1
GENERAL INFORMATION:
APPLICANT: Alderson, Mark
APPLICANT: Dillon, Davin C.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Campos-Neto, Antonio
TITLE OF INVENTION: COMPOUNDS AND DIAGNOSIS OF
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY
STREET: 6300 Columbia Center, 701 Fifth Ave.
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,009
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.441C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
US-09-073-009-75

Query Match 2.8%; Score 6; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 183 SDVLTA 188
Db 3 SDVLTA 8

RESULT 59
US-09-793-306-73
Sequence 73, Application US/09793306
Patent No. US20020098200A1
GENERAL INFORMATION:
APPLICANT: Campos-Neto, Antonio
APPLICANT: Skeiky, Yasir
APPLICANT: Ovendale, Pamela
APPLICANT: Jen, Shyian
APPLICANT: Lodes, Michael
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compounds and Methods for Diagnosis and Immunotherapy
TITLE OF INVENTION: of Tuberculosis
FILE REFERENCE: 014058-008740US
CURRENT APPLICATION NUMBER: US/09/793,306
CURRENT FILING DATE: 2001-02-26
PRIOR APPLICATION NUMBER: US 60/185,037
PRIOR FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: US 60/223,828
PRIOR FILING DATE: 2000-08-08
NUMBER OF SEQ ID NOS: 164
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 73
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:peptide
OTHER INFORMATION: ORF-2-23
US-09-793-306-73

Query Match 2.8%; Score 6; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 183 SDVLTA 188
Db 8 SDVLTA 13

RESULT 60
US-09-793-306-75
Sequence 75, Application US/09793306
Patent No. US20020098200A1
GENERAL INFORMATION:
APPLICANT: Campos-Neto, Antonio
APPLICANT: Skeiky, Yasir
APPLICANT: Ovendale, Pamela
APPLICANT: Jen, Shyian
APPLICANT: Lodes, Michael
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compounds and Methods for Diagnosis and Immunotherapy
TITLE OF INVENTION: of Tuberculosis
FILE REFERENCE: 014058-008740US
CURRENT APPLICATION NUMBER: US/09/793,306
CURRENT FILING DATE: 2001-02-26
PRIOR APPLICATION NUMBER: US 60/185,037

;; PRIOR FILING DATE: 2000-02-25
;; PRIOR APPLICATION NUMBER: US 60/223,828
;; PRIOR FILING DATE: 2000-08-08
;; NUMBER OF SEQ ID NOS: 164
;; SOFTWARE: Patent In Ver. 2.1
;; SEQ ID NO 75
;; LENGTH: 15
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence:peptide
;; OTHER INFORMATION: ORF-2-25
US-09-793-306-75

Query Match 2.8%; Score 6; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 183 SDVLTA 188
Db 3 SDVLTA 8

RESULT 61
US-10-215-435-3
;; Sequence 3, Application US/10215435
;; Publication No. US20030104614A1
;; GENERAL INFORMATION:
;; APPLICANT: Uhrich, Kathryn E.
;; APPLICANT: Buettner, Helen
;; APPLICANT: Schmalenberg, Kristine
;; TITLE OF INVENTION: Micro patterning Surfaces of Polymeric Substrates
;; FILE REFERENCE: 1435.009US1
;; CURRENT APPLICATION NUMBER: US/10/215,435
;; CURRENT FILING DATE: 2002-08-09
;; PRIOR APPLICATION NUMBER: PCT/US01/04842
;; PRIOR FILING DATE: 2001-02-12
;; PRIOR APPLICATION NUMBER: US 60/181,763
;; PRIOR FILING DATE: 2000-02-11
;; NUMBER OF SEQ ID NOS: 13
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 3
;; LENGTH: 19
;; TYPE: PRT
;; ORGANISM: Unknown
;; FEATURE:
;; OTHER INFORMATION: A peptide sequence found in the A chain of laminin.
US-10-215-435-3

Query Match 2.8%; Score 6; DB 14; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 VAVSAD 18
Db 13 VAVSAD 18

RESULT 62
US-10-622-072-3
;; Sequence 3, Application US/10622072
;; Publication No. US20040096478A1
;; GENERAL INFORMATION:
;; APPLICANT: Uhrich, K. E.
;; APPLICANT: Schmalenberg, K.
;; TITLE OF INVENTION: Therapeutic devices for patterned cell growth
;; FILE REFERENCE: 1435.028US1
;; CURRENT APPLICATION NUMBER: US/10/622,072
;; CURRENT FILING DATE: 2003-07-17
;; PRIOR APPLICATION NUMBER: US 60/396,628
;; PRIOR FILING DATE: 2002-07-17
;; NUMBER OF SEQ ID NOS: 13
;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 3
;; LENGTH: 19
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: A water-soluble peptide.
US-10-622-072-3

Query Match 2.8%; Score 6; DB 16; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 VAVSAD 18
Db 13 VAVSAD 18

RESULT 63
US-09-864-761-36329
;; Sequence 36329, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;; FILE REFERENCE: Aseomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 36329
;; LENGTH: 21
;; TYPE: PRT
;; ORGANISM: Homo sapiens

FEATURE:
; OTHER INFORMATION: MAP TO AC008015.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 9.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.2
; OTHER INFORMATION: EST_HUMAN HIT: AW874553.1, EVALUATE 2.00e-04
US-09-864-761-36329

Query Match 2.8%; Score 6; DB 9; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 TKVLAP 204
Db 10 TKVLAP 15

RESULT 64
US-10-360-053-23
; Sequence 23, Application US/10360053
; Publication No. US20030170230A1
; GENERAL INFORMATION:
; APPLICANT: Caterer, Nigel
; APPLICANT: Utterthal, Lars O
; APPLICANT: Nielsen, Rasmus W
; TITLE OF INVENTION: Compositions and Methods for Assembly and Stabilization of Antibody
; TITLE OF INVENTION: Fragments via Antiparallel Heterogeneous Coiled-Coil Peptide Residues
; FILE REFERENCE: IMX-0028
; CURRENT APPLICATION NUMBER: US/10/360,053
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US60/354,376
; PRIOR FILING DATE: 2002-02-05
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-360-053-23

Query Match 2.8%; Score 6; DB 14; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 IAAIEQ 42
Db 20 IAAIEQ 25

RESULT 65
US-10-424-599-254894
; Sequence 254894, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kowalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21/53223/B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 254894
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_72192C.1.pap
US-10-424-599-254894

Query Match 2.8%; Score 6; DB 12; Length 32;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 FGVLSL 184
Db 3 FGVLSL 8

RESULT 66
US-09-864-761-41249
; Sequence 41249, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41249
; LENGTH: 33
; TYPE: PRT

; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC012308.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.82
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EST HUMAN HIT: BF082286.1, EVALUE 8.00e-12
; OTHER INFORMATION: SWISSPROT HIT: P07664, EVALUE 6.50e+00
US-09-864-761-41249

Query Match 2.8%; Score 6; DB 9; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 ITEIN 30
Db 12 ITEIN 17
|||||

RESULT 67
US-10-424-599-172112
; Sequence 172112, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; SEQ ID NO 172112
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_126430C.1.pep
US-10-424-599-172112

Query Match 2.8%; Score 6; DB 12; Length 37;
Best Local Similarity 100.0%; Pred. No. 3.5e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 DPIHYD 23
Db 15 DPIHYD 20
|||||

RESULT 68
US-10-424-599-245873
; Sequence 245873, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 245873
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL157773.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.1

; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_64054C.1.pep
US-10-424-599-245873

Query Match 2.8%; Score 6; DB 12; Length 42;
Best Local Similarity 100.0%; Pred. No. 3.9e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 186 LTAIFQ 191
Db 12 LTAIFQ 17
|||||

RESULT 69
US-09-864-761-40280
; Sequence 40280, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40280
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL157773.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.1

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.4
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.8
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.8
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 6.4
US-09-864-761-40280

Query Match 2.8%; Score 6; DB 9; Length 43;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LIAAVA 11
Db 5 LIAAVA 10
|||||

RESULT 70

US-10-424-599-219711
; Sequence 219711, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 219711
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_40427C.1.pep
US-10-424-599-219711

Query Match 2.8%; Score 6; DB 12; Length 43;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 TVRKEM 198
Db 24 TVRKEM 29
|||||

RESULT 71

US-10-424-599-257467
; Sequence 257467, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 257467
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_74516C.1.pep
US-10-424-599-257467

Query Match 2.8%; Score 6; DB 12; Length 45;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSAD 18
Db 29 VAVSAD 34
|||||

RESULT 72

US-10-424-599-188617
; Sequence 188617, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 188617
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_141335C.1.pep
US-10-424-599-188617

Query Match 2.8%; Score 6; DB 12; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 TTHVIS 133
Db 44 TTHVIS 49
|||||

RESULT 73

US-10-424-599-182164
; Sequence 182164, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 182164
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_135507C.1.pep
US-10-424-599-182164

Query Match 2.8%; Score 6; DB 12; Length 51;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 BEINKA 32
Db 30 BEINKA 35
|||||

```
RESULT 74
US-10-437-963-142935
; Sequence 142935, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 142935
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_43894C.1.pep
US-10-437-963-142935

Query Match      2.8%; Score 6; DB 16; Length 53;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 RKEMTK 200
DB 36 RKEMTK 41

RESULT 75
US-10-424-599-148141
; Sequence 148141, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 148141
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_104794C.1.pep
US-10-424-599-148141

Query Match      2.8%; Score 6; DB 12; Length 54;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 EISDEG 150
DB 4 EISDEG 9

RESULT 76
US-10-424-599-266220
; Sequence 266220, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 142935
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_43894C.1.pep
US-10-437-963-142935

Query Match      2.8%; Score 6; DB 16; Length 53;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 RKEMTK 200
DB 36 RKEMTK 41

RESULT 77
US-10-424-599-236065
; Sequence 236065, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 236065
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_55195C.1.pep
US-10-424-599-236065

Query Match      2.8%; Score 6; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 TKVLAP 204
DB 33 TKVLAP 38

RESULT 78
US-10-264-049-3669
; Sequence 3669, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
US-10-264-049-3669
```

; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3669
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-3669

Query Match 2.8%; Score 6; DB 15; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 HDDIVS 113
Db 45 HDDIVS 50

RESULT 79
US-10-211-962-124
; Sequence 124, Application US/10211962
; Publication No. US20030082640A1
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/10/211,962
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US/09/562,737
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 124
; LENGTH: 57
; TYPE: PRT
; ORGANISM: mouse
US-10-211-962-124

Query Match 2.8%; Score 6; DB 14; Length 57;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 VALSLE 145
Db 47 VALSLE 52

RESULT 80
US-10-437-963-188318
; Sequence 188318, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 188318
; LENGTH: 57
; TYPE: PRT

; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) (57)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MET4530_84934C.1.pap
US-10-437-963-188318

Query Match 2.8%; Score 6; DB 16; Length 57;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 DATAI 40
Db 11 DATAI 16

RESULT 81
US-09-864-761-35621
; Sequence 35621, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35621
; LENGTH: 58
; TYPE: PRT

; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007089.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EST_HUMAN HIT: BF129244.1, EVALUE 1.20e+00
US-09-864-761-35621

Query Match 2.8%; Score 6; DB 9; Length 59;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 AVSADP 19
Db 39 AVSADP 44
|||||

RESULT 82

US-10-424-599-143933
; Sequence 143933, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 143933
; LENGTH: 59
; TYPE: PRT

; ORGANISM: Glycine max
; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_100985C.1.pep
US-10-424-599-143933

Query Match 2.8%; Score 6; DB 12; Length 59;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 KFLIIA 8
Db 36 KFLIIA 41
|||||

RESULT 83

US-10-424-599-252348
; Sequence 252348, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 252348

; LENGTH: 59
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_6989C.1.pep
US-10-424-599-252348

Query Match 2.8%; Score 6; DB 12; Length 59;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 93 KGEEGI 98
Db 54 KGEEGI 59
|||||

RESULT 84

US-10-424-599-245904
; Sequence 245904, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 245904
; LENGTH: 60
; TYPE: PRT

; ORGANISM: Glycine max
; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_64082C.1.pep
US-10-424-599-245904

Query Match 2.8%; Score 6; DB 12; Length 60;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VSADPI 20
Db 13 VSADPI 18
|||||

RESULT 85

US-10-437-963-197055
; Sequence 197055, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 197055
; LENGTH: 61
; TYPE: PRT

; ORGANISM: Oryza sativa
; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_9284C.1.pep

```
US-10-437-963-197055
Query Match      2.8%; Score 6; DB 16; Length 61;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 AFVAVS 16
    |||||
Db 41 AFVAVS 46

RESULT 86
US-10-424-599-285188
; Sequence 285188, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 285188
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_99554C.1.pep
US-10-424-599-285188

Query Match      2.8%; Score 6; DB 12; Length 64;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GVLSDV 185
    |||||
Db 45 GVLSDV 50

RESULT 87
US-10-424-599-244365
; Sequence 244365, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 244365
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_62689C.1.pep
US-10-424-599-244365

Query Match      2.8%; Score 6; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 GGLSIL 175
    |||||
Db 33 GGLSIL 38

US-10-437-963-108901
; Sequence 108901, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 108901
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_1310C.1.pep
US-10-437-963-108901

Query Match      2.8%; Score 6; DB 16; Length 65;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 202 LAPAFK 207
    |||||
Db 8 LAPAFK 13

RESULT 89
US-10-437-963-119979
; Sequence 119979, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 119979
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_23142C.1.pep
US-10-437-963-119979

Query Match      2.8%; Score 6; DB 16; Length 65;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVALSL 144
    |||||
Db 36 VVALSL 41
```

```

RESULT 90
US-10-437-963-143189
; Sequence 143189, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 143189
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_44121C.1.pep
US-10-437-963-143189

Query Match      2.8%; Score 6; DB 16; Length 65;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 HIGGLS 173
DB 58 HIGGLS 63

RESULT 91
US-10-424-599-228585
; Sequence 228585, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 228585
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_4843C.1.pep
US-10-424-599-228585

Query Match      2.8%; Score 6; DB 12; Length 66;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 KAHLLI 105
DB 35 KAHLLI 40

RESULT 92
US-10-437-963-163372
; Sequence 163372, Application US/10437963
; Publication No. US20040123343A1

```

```

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 163372
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_62375C.1.pep
US-10-437-963-163372

Query Match      2.8%; Score 6; DB 16; Length 66;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AAVAFV 13
DB 49 AAVAFV 54

RESULT 93
US-10-437-963-184690
; Sequence 184690, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 184690
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(66)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_8165C.1.pep
US-10-437-963-184690

Query Match      2.8%; Score 6; DB 16; Length 66;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 AAVAFV 14
DB 4 AAVAFV 9

RESULT 94

```

US-10-424-599-179217
; Sequence 179217, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 179217
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_13284C.1.pep
US-10-424-599-179217

Query Match 2.8%; Score 6; DB 12; Length 67;
Best Local Similarity 100.0%; Pred.No. 6e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 LIGVHD 109
| | | | |
DB 29 LIGVHD 34

RESULT 95
US-09-738-626-6222
; Sequence 6222, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6222
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6222

Query Match 2.8%; Score 6; DB 9; Length 69;
Best Local Similarity 100.0%; Pred.No. 6.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PLLTAA 9
| | | | |
DB 17 PLLTAA 22

RESULT 96
US-10-424-599-181866
; Sequence 181866, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 181866
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_135239C.1.pep
US-10-424-599-181866

Query Match 2.8%; Score 6; DB 12; Length 69;
Best Local Similarity 100.0%; Pred.No. 6.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 HDDIVS 113
| | | | |
DB 42 HDDIVS 47

RESULT 97
US-10-424-599-227078
; Sequence 227078, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 227078
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_47081C.1.pep
US-10-424-599-227078

Query Match 2.8%; Score 6; DB 12; Length 72;
Best Local Similarity 100.0%; Pred.No. 6.4e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 FGVLSD 184
| | | | |
DB 16 FGVLSD 21

RESULT 98
US-10-424-599-282630
; Sequence 282630, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 282630
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_97236C.1.pep
US-10-424-599-282630

Query Match 2.8%; Score 6; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 IFQDTV 194
Db 23 IFQDTV 28

RESULT 99

US-10-424-599-151004
; Sequence 151004, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 151004
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(74)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_107380C.1.pep
US-10-424-599-151004

Query Match 2.8%; Score 6; DB 12; Length 74;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 IEQSET 45
Db 21 IEQSET 26

RESULT 100

US-10-425-114-40082
; Sequence 40082, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 40082
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700989818_FLI.pep
US-10-425-114-40082

Query Match 2.8%; Score 6; DB 12; Length 76;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 LAYKLG 123
Db 15 LAYKLG 20

Search completed: August 6, 2004, 16:15:10
Job time : 47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2004, 16:06:28 ; Search time 17 Seconds
(without alignments)
1205.223 Million cell updates/sec

Title: US-10-024-955-7

Perfect score: 213

Sequence: 1 MMKFLJJAAVAVASADPI.....VRXEMTKVLAPAFKRELEKN 213

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : PIR_78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.8	349	2 T06680	hypothetical prote
2	8	3.8	528	2 JN0445	transcription init
3	8	3.8	536	1 F0WVMD	gag polyprotein -
4	8	3.8	783	2 T01015	probable subtilisi
5	8	3.8	918	2 T02759	hypothetical prote
6	8	3.8	1784	2 T10332	gag-pol polyprotei
7	7	3.3	60	2 PN0674	GRP-binding protei
8	7	3.3	72	2 A45009	rho protein GDP-di
9	7	3.3	87	1 A44562	phosphotransferase
10	7	3.3	87	2 E95136	phosphocarrier pro
11	7	3.3	87	2 G98004	hypothetical prote
12	7	3.3	106	2 S38768	rho protein GDP di
13	7	3.3	140	2 H71197	hypothetical prote
14	7	3.3	149	2 D69401	hypothetical prote
15	7	3.3	151	2 E84319	hypothetical prote
16	7	3.3	162	2 T10831	hypothetical prote
17	7	3.3	169	2 S03744	Y4xB protein - Rhi
18	7	3.3	183	2 T51495	hypothetical prote
19	7	3.3	185	2 G70425	hypothetical prote
20	7	3.3	190	2 AC3049	hypothetical prote
21	7	3.3	190	2 H98236	ornithine cyclodea
22	7	3.3	202	2 B90441	hypothetical prote
23	7	3.3	204	2 S12121	hypothetical prote
24	7	3.3	204	2 I38156	rho protein GDP-di
25	7	3.3	208	2 T27749	rho protein GDP-di
26	7	3.3	251	2 T42005	hypothetical prote
27	7	3.3	252	2 H82219	conserved hypothet
28	7	3.3	253	2 G64320	conserved hypothet
29	7	3.3	286	2 A90246	hypothetical prote

30	7	3.3	295	2 S29956	beta-chimerin - ra
31	7	3.3	300	2 F71616	RNA-binding protei
32	7	3.3	305	2 H87124	hypothetical prote
33	7	3.3	307	2 T46938	ornithine cyclodea
34	7	3.3	310	2 A02243	hypothetical prote
35	7	3.3	333	2 A72371	oligopeptide ABC t
36	7	3.3	345	1 G97024	probable phosphoe
37	7	3.3	349	2 T17130	hypothetical prote
38	7	3.3	354	1 DUAGO	ornithine cyclodea
39	7	3.3	354	2 S55582	ornithine cyclodea
40	7	3.3	356	2 A32049	ornithine cyclodea
41	7	3.3	356	2 B95993	probable iron ABC
42	7	3.3	371	2 AF3229	ornithine cyclodea
43	7	3.3	375	2 T15356	hypothetical prote
44	7	3.3	408	2 E84743	hypothetical prote
45	7	3.3	415	2 D96664	hypothetical prote
46	7	3.3	427	2 T17123	hypothetical prote
47	7	3.3	430	2 C97237	histidyl-tRNA synt
48	7	3.3	431	2 C86178	hypothetical prote
49	7	3.3	434	2 C86768	xanthine permease
50	7	3.3	443	2 B53764	beta2-chimerin, ce
51	7	3.3	466	2 A53764	beta2-chimerin, ce
52	7	3.3	474	2 T01347	pectinesterase hom
53	7	3.3	500	2 A80274	L-arabinose isomer
54	7	3.3	556	2 T40077	hypothetical prote
55	7	3.3	572	2 H85731	Rhs element associ
56	7	3.3	573	2 A53035	conserved hypothet
57	7	3.3	573	2 F98250	hypothetical prote
58	7	3.3	595	2 E87698	sensor histidine k
59	7	3.3	628	2 B87448	conserved hypothet
60	7	3.3	657	2 E70529	hypothetical prote
61	7	3.3	736	2 H75460	conserved hypothet
62	7	3.3	829	2 S12858	virA protein - Agr
63	7	3.3	833	2 S06972	virA protein - Agr
64	7	3.3	833	2 A32448	two component sens
65	7	3.3	871	2 H72597	hypothetical prote
66	7	3.3	996	2 C75498	probable metallopr
67	7	3.3	1027	2 H90861	probable multidrug
68	7	3.3	1027	2 D85757	probable efflux pu
69	7	3.3	1155	2 B71720	hypothetical prote
70	7	3.3	1194	1 G70837	probable ABC trans
71	7	3.3	1398	2 H90698	Rhs core protein
72	7	3.3	1398	2 B85549	hypothetical prote
73	7	3.3	1400	2 E30886	RhsE core protein
74	7	3.3	2009	2 S49764	SBC7 protein - Vea
75	7	3.3	2033	2 T30849	actin binding prot
76	6	2.8	34	2 S71457	NADH2 dehydrogenas
77	6	2.8	34	2 G84147	hypothetical prote
78	6	2.8	40	2 S23286	light-harvesting p
79	6	2.8	48	2 S68885	light-harvesting p
80	6	2.8	48	2 S68886	light-harvesting p
81	6	2.8	49	2 S55781	laminin alpha chai
82	6	2.8	52	2 A37354	sex-specific prote
83	6	2.8	62	2 G30111	hypothetical prote
84	6	2.8	72	2 D84194	hypothetical prote
85	6	2.8	75	2 F75031	hydrogenase expres
86	6	2.8	76	2 E69472	conserved hypothet
87	6	2.8	77	2 B83961	acyl-carrier prote
88	6	2.8	77	2 H69420	hydrogenase expres
89	6	2.8	80	2 E83102	conserved hypothet
90	6	2.8	80	2 C82809	hypothetical prote
91	6	2.8	82	2 A72125	hypothetical prote
92	6	2.8	82	2 S73167	H+-transporting tw
93	6	2.8	82	2 S26958	H+-transporting tw
94	6	2.8	83	2 S39516	H+-transporting tw
95	6	2.8	83	2 S23424	H+-transporting tw
96	6	2.8	85	2 D83389	conserved hypothet
97	6	2.8	87	2 H64540	hypothetical prote
98	6	2.8	87	2 C71966	hypothetical prote
99	6	2.8	92	2 C90667	hypothetical membr
100	6	2.8	92	2 G85517	hypothetical prote

ALIGNMENTS

RESULT 1

T06680. hypothetical protein T17F15.100 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 02-Sep-2000
C:Accession: T06680
R:Quekter, F.; Choinsne, N.; Robert, C.; Brottier, P.; Cattolico, L.; Artigou
submitted to the Protein Sequence Database, April 1999
A:Reference number: Z15793
A:Accession: T06680
A:Molecule type: DNA
A:Residues: 1-349 <QUE>
A:Cross-references: EMBL:AL049658; GSPDB:GN00061; ATSP:T17F15.100
A:Experimental source: cultivar Columbia; BAC clone T17F15
C:Genetics:
A:Gene: ATSP:T17F15.100
A:Map position: 3
A:Introns: 26/1; 81/2
C:Superfamily: RING finger homology
P:203-254/Domain: RING finger homology <ERN>

C/Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 08-Oct-1999
C/Accession: T02759
R/Segura, A.; Ornsdon, N.L.
Submitted to the EMBL Data Library, June 1997
A/Description: PZRI53.
A/Reference number: Z14724
A/Accession: T02759
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-918 <SEG>
A/Cross-references: EMBL:AF011339; NID:G2286203; PIDN:AAC27114.1; PID:G2286204
A/Experimental source: strain ADP1

Query Match 3.8%; Score 8; DB 2; Length 918;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 INKAIDDA 36
|||||||
Db 78 INKAIDDA 85

RESULT 6
T10532
gag-pol polyprotein precursor - feline leukemia virus (strain FELV-FAIDS)
N/Contains: core protein p10; core protein p12; core protein p15; core protein p27; endo
C/Species: feline leukemia virus
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C/Accession: T10532
R/Donahue, P.R.; Hoover, E.A.; Beltz, G.A.; Riedel, N.; Hirsch, V.M.; Overbaugh, J.; Mul
J. Virol. 62, 722-731, 1988
A/Title: Strong sequence conservation among horizontally transmissible, minimally pathog
A/Reference number: Z17078; MUID:88119207; PMID:2828667
A/Accession: T10532
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1784 <DON>
A/Cross-references: EMBL:M18247; NID:G323904; PIDN:AAA93092.1; PID:G1237256
C/Suprafamily: pol polyprotein
C/Keywords: hydrolase; nucleotidyltransferase; polyprotein; reverse transcriptase
F/75-576/Product: gag polyprotein #status predicted <GAG>
F/75-201/Product: core protein p15 #status predicted <P15>
F/202-271/Product: core protein p12 #status predicted <P12>
F/272-519/Product: core protein p27 #status predicted <P27>
F/520-576/Product: core protein p10 #status predicted <P10>
F/577-1784/Product: pol polyprotein #status predicted <POL>
F/577-701/Product: proteinase #status predicted <PTN>
F/702-1368/Product: RNA-directed DNA polymerase (EC 2.7.7.49) #status predicted <REV>
F/1369-1784/Product: endonuclease (EC 3.1.-.-) #status predicted <EDC>

Query Match 3.8%; Score 8; DB 2; Length 1784;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 KEMTKVLA 203
|||||||
Db 513 KEMTKVLA 520

RESULT 7
PN0674
GTP-binding protein G25K - human (fragments)
N/Alternate names: guanine nucleotide dissociation inhibitor protein
C/Species: Homo sapiens (man)
C/Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 18-Aug-2000
C/Accession: PN0674
R/Backlund Jr., P.S.
Biochem. Biophys. Res. Commun. 196, 534-542, 1993
A/Title: Carboxyl methylation of the low molecular weight GTP-binding protein G25K: Regu
A/Reference number: PN0674; MUID:94059070; PMID:8240325
A/Accession: PN0674
A/Molecule type: protein
A/Residues: 1-60 <BAC>

A/Experimental source: brain
C/Comment: This protein is a regulator of G25K methylation by blocking methylation of G2E
C/Suprafamily: human rho GDP dissociation inhibitor
C/Keywords: GTP binding

Query Match 3.3%; Score 7; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSADP 19
|||||||
Db 7 VAVSADP 13

RESULT 8
A45009
rho protein GDP-dissociation inhibitor - bovine (fragments)
C/Species: Bos primigenius taurus (cattle)
C/Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 18-Aug-2000
C/Accession: A45009
R/Bourmeyster, N.; Stasia, M.J.; Garin, J.; Gagnon, J.; Boquet, P.; Vignais, P.V.
Biochemistry 31, 12863-12869, 1992
A/Title: Copurification of rho protein and the rho-GDP dissociation inhibitor from bovine
tulinum.
A/Reference number: A45009; MUID:93099151; PMID:1334435
A/Accession: A45009
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-72 <BOU>
C/Suprafamily: human rho GDP dissociation inhibitor

Query Match 3.3%; Score 7; DB 2; Length 72;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSADP 19
|||||||
Db 17 VAVSADP 23

RESULT 9
A44562
phosphotransferase system phosphohistidine-containing protein - Streptococcus mutans
N/Alternate names: phosphocarrier protein HPr
C/Species: Streptococcus mutans
C/Date: 10-Mar-1994 #sequence_revision 31-Jan-1997 #text_change 26-Feb-1999
C/Accession: A44562; JC2114
R/Reynolds, B.C.
Submitted to the Protein Sequence Database, February 1994
A/Reference number: A44562
A/Accession: A44562
A/Molecule type: protein
A/Residues: 1-87 <REV>
A/Experimental source: strain Ingbritt
A/Note: the initial Met was found in 14% of the molecules
R/Dashper, S.G.; Kirsbaum, L.; Hug, N.L.; Riley, P.F.; Reynolds, E.C.
Biochem. Biophys. Res. Commun. 199, 1297-1304, 1994
A/Title: Complete amino acid sequence and comparative molecular modelling of HPr from Str
A/Reference number: JC2114; MUID:94197719; PMID:8147873
A/Accession: JC2114
A/Molecule type: protein
A/Residues: 2-87 <BAS>
A/Note: Initiator Met was present in 14% of the molecules
C/Comment: This protein is a heat stable regulatory protein of the phosphotransferase sys
C/Suprafamily: phosphocarrier protein; phosphohistidine-containing protein; phosphotran
C/Keywords: phosphocarrier protein; phosphohistidine; phosphoprotein; sugar transport sy
F/8-85/Domain: phosphotransferase system phosphohistidine-containing protein homology <H
F/1/Modified site: N-formylmethionine (partial) #status experimental
F/15/Binding site: phosphate (His) (covalent) #status experimental
F/46/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 3.3%; Score 7; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 20;

Matches	7;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	34	DDAIAAI	40						
Db	71	DDAIAAI	77						
RESULT 10									
E95136									
phosphocarrier protein HPr [imported] - Streptococcus pneumoniae (strain TIGR4)									
C:Species: Streptococcus pneumoniae									
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 17-May-2002									
C:Accession: E95136									
R:Titelstein, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid									
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,									
nson, T.; Hickey, E.K.; Holt, I.E.									
Science 293, 498-506, 2001									
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,									
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.									
A:Reference number: A95000; MUID:21357209; PMID:11463916									
A:Accession: E95136									
A:Status: preliminary									
A:Molecule type: DNA									
A:Residues: 1-87 <KUR>									
A:Cross-references: GB:AE005672; PIDN:AAK75286.1; PID:gl4972657; GSPDB:GN00164; TIGR:SP4									
A:Experimental source: strain TIGR4									
C:Genetics:									
A:Gene: SP1177									
C:Superfamily: phosphotransferase system phosphohistidine-containing protein; phosphotra									
Query Match 3.3%; Score 7; DB 2; Length 87;									
Best Local Similarity 100.0%; Pred. No. 20;									
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	34	DDAIAAI	40						
Db	71	DDAIAAI	77						
RESULT 11									
G98004									
hypothetical protein p7SH [imported] - Streptococcus pneumoniae (strain R6)									
C:Species: Streptococcus pneumoniae									
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 17-May-2002									
C:Accession: G98004									
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E									
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M									
Y, P.; Sun, P.M.; Winkler, M.E.									
J. Bacteriol. 183, 5709-5717, 2001									
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;									
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.									
A:Reference number: A97872; MUID:21429245; PMID:11544234									
A:Accession: G98004									
A:Status: preliminary									
A:Molecule type: DNA									
A:Residues: 1-87 <KUR>									
A:Cross-references: GB:AE007317; PIDN:AAK99867.1; PID:gl5458685; GSPDB:GN00174									
C:Genetics:									
A:Gene: p7SH									
C:Superfamily: phosphotransferase system phosphohistidine-containing protein; phosphotra									
Query Match 3.3%; Score 7; DB 2; Length 87;									
Best Local Similarity 100.0%; Pred. No. 20;									
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	34	DDAIAAI	40						
Db	71	DDAIAAI	77						
RESULT 12									
S38768									
rho protein GDP dissociation inhibitor - guinea pig (fragments)									

C:Species: Cavia porcellus (guinea pig)									
C:Date: 09-Jun-1994 #sequence_revision 06-Sep-1996 #text_change 18-Aug-2000									
C:Accession: S38768									
R:Pick, E.; Gorzalczany, Y.; Engel, S.									
Eur. J. Biochem. 217, 441-455, 1993									
A:Title: Role of the rac1 p21-GDP-dissociation inhibitor for rho heterodimer in the activ									
A:Reference number: S38767; MUID:94039069; PMID:8223583									
A:Accession: S38768									
A:Molecule type: protein									
A:Residues: 1-40;41-46;47-86;87-106 <PIC>									
C:Superfamily: human rho GDP dissociation inhibitor									
Query Match 3.3%; Score 7; DB 2; Length 106;									
Best Local Similarity 100.0%; Pred. No. 24;									
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	13	VAVSADP	19						
Db	1	VAVSADP	7						
RESULT 13									
H71197									
hypothetical protein PH1854 - Pyrococcus horikoshii									
C:Species: Pyrococcus horikoshii									
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 28-Jul-2000									
C:Accession: H71197									
R:Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekine									
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi,									
DNA Res. 5, 55-76, 1998									
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic an									
A:Reference number: A71000; MUID:98344137; PMID:9679194									
A:Accession: H71197									
A:Status: preliminary; nucleic acid sequence not shown; translation not shown									
A:Molecule type: DNA									
A:Residues: 1-140 <KAW>									
A:Cross-references: GB:AP000007; NID:g3236134; PIDN:BAA30975.1; PID:g3258292									
A:Experimental source: strain O13									
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank									
C:Genetics:									
A:Gene: PH1854									
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1854									
Query Match 3.3%; Score 7; DB 2; Length 140;									
Best Local Similarity 100.0%; Pred. No. 30;									
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	25	ITEINK	31						
Db	8	ITEINK	14						
RESULT 14									
D69401									
hypothetical protein AF1213 - Archaeoglobus fulgidus									
C:Species: Archaeoglobus fulgidus									
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999									
C:Accession: D69401									
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,									
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;									
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.									
Nature 390, 364-370, 1997									
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes, S.N.;									
Smith, H.O.; Woese, C.R.; Venter, J.C.									
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo									
A:Reference number: A69250; MUID:98049343; PMID:9389475									
A:Accession: D69401									
A:Status: preliminary; nucleic acid sequence not shown; translation not shown									
A:Molecule type: DNA									
A:Residues: 1-149 <KLE>									
A:Cross-references: GB:AE001020; GB:AE000782; NID:g2699343; PIDN:AAB90034.1; PID:g264937									
Query Match 3.3%; Score 7; DB 2; Length 149;									

Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 FKRELEK 212
Db 7 FKRELEK 13

RESULT 15
E84319
hypothetical protein Vngl670c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: E84319
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Beckwith, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablonka, K.H.; Alam, M.; Freitas, I.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li, A.; Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: E84319
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-151 <STO>
A;Cross-references: GB:AE004437; NID:gl0581141; PIDN:AAG19921.1; GSPDB:GN00138
C;Genetics:
A;Gene: VNG1670C

Query Match 3.3%; Score 7; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 FVVALSL 144
Db 24 FVVALSL 30

RESULT 16
Tl0831
Y4XD protein - Rhizobium sp. (strain NGR234) plasmid pNGR234a
C;Species: Rhizobium sp.
A;Variety: strain NGR234
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 02-Sep-2000
C;Accession: Tl0831
R;Freiberg, C.; Fellay, R.; Bairoch, A.; Broughton, W.J.; Rosenthal, A.; Perret, X.
Nature 387, 394-401, 1997
A;Title: Molecular basis of symbiosis between Rhizobium and legumes.
A;Reference number: Z14734; MUID:97305956; PMID:9163424
A;Accession: Tl0831
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-162 <PRE>
A;Cross-references: EMBL:AE000105; NID:G2182706; PID:G2182710
C;Genetics:
A;Gene: Y4XD
A;Genome: plasmid pNGR234a
C;Superfamily: Rhodobacter hypothetical 17.5K protein (nifx 3' region)

Query Match 3.3%; Score 7; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 LEISDEG 150
Db 100 LEISDEG 106

RESULT 17
S03744
hypothetical protein - Calothrix sp.
C;Species: Calothrix sp.
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 18-Jun-1993

C;Accession: S03744
R;Parsot, C.; Mazel, D.
Mol. Microbiol. 1, 45-52, 1987
A;Title: Cloning and nucleotide sequence of the thrB gene from the cyanobacterium Calothrix
A;Reference number: S03743; MUID:88260883; PMID:2838727
A;Accession: S03744
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-169 <PAR>

Query Match 3.3%; Score 7; DB 2; Length 169;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 IDDAIAA 39
Db 62 IDDAIAA 68

RESULT 18
T51495
hypothetical protein T21H19_170 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
C;Accession: T51495
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51495
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-183 <SAT>
A;Cross-references: EMBL:AL391148
A;Experimental source: cultivar Columbia; BAC clone T21H19
C;Genetics:
A;Map position: 5
A;Note: T21H19_170

Query Match 3.3%; Score 7; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 185 VLTAIFQ 191
Db 87 VLTAIFQ 93

RESULT 19
G70425
hypothetical protein aq_1446 - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C;Accession: G70425
R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Over
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: G70425
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-185 <AQF>
A;Cross-references: GB:AE000741; NID:G2983841; PIDN:AA07402.1; PID:G2983848; GB:AE00065;
A;Experimental source: strain VF5
C;Genetics:
A;Gene: aq_1446

Query Match 3.3%; Score 7; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMKPLLI 7
|||||


```

Query Match          3.3%; Score 7; DB 2; Length 251;
Best Local Similarity 100.0%; Pred.No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      180 GVLSDVL 186
      |||||
Db       82 GVLSDVL 88

RESULT 27
H82219 conserved hypothetical protein VC1285 [imported] - Vibrio cholerae (strain N16961 serog
C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C/Accession: H82219
R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A/Reference number: A82035; MUID:20406833; PMID:10952301
A/Accession: H82219
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-252 <HEI>
A/Cross-references: GB:AE004207; GS:AE003852; NID:g9655761; PIDN:AAF94444.1; GSPDB:GN0001

```

Db 259 FKRELEK 265

```
RESULT 30
S29956
Beta-Chimerin - rat
N;Alternate names: GTPase-activating protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C;Accession: A45485; S31398; S29956
R;Leung, T.; How, B.E.; Manser, E.; Lim, L.
J. Biol. Chem. 268, 3813-3816, 1993
A;Title: Germ cell beta-chimaerin, a new GTPase-activating protein for p21rac, is specific
A;Reference number: A45485; MUID:93179371; PMID:8440677
A;Accession: A45485
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-295 <LEU>
A;Cross-references: EMBL:X69489; NID:g57526; PIDN:CAA49244.1; PID:g57527
A;Experimental source: testis
A;Note: sequence extracted from NCBI backbone (NCBIP:125731)
R;Leung, T.; How, B.E.; Manser, E.; Lim, L.
submitted to the EMBL Data Library, November 1992
A;Description: Germ cell beta-chimaerin, a new GTPase-activating protein for p21rac, is
A;Reference number: S31398
A;Accession: S31398
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-295 <LE2>
A;Cross-references: EMBL:X69462; NID:g55816; PID:g55817
C;Superfamily: protein kinase C zinc-binding repeat homology
F;42-91/Domain: protein kinase C zinc-binding repeat homology <kz2>

Query Match 3.3%; Score 7; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 IEARGLK 82
Db 128 IEARGLK 134

RESULT 31
F71616
RNA-binding protein (KH domain) PFB0370c - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 08-Oct-1999
C;Accession: F71616
R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
; Partea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A;Reference number: A71600; MUID:99021743; PMID:9804551
A;Accession: F71616
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-300 <GAR>
A;Cross-references: GB:AE001390; GB:AE001362; NID:g3845164; PIDN:AAC71863.1; PID:g384516
A;Experimental source: clone 3D7
C;Genetics:
A;Gene: PFB0370c
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0443

Query Match 3.3%; Score 7; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 207 KRELEKN 213
Db 214 KRELEKN 220

RESULT 32
H87124
hypothetical protein ML1726 [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 24-Aug-2001
C;Accession: H87124
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Hor
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutherford, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: A86909; MUID:21128732; PMID:11234002
A;Accession: H87124
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-305 <STO>
A;Cross-references: GB:AL450380; NID:gl3093478; PIDN:CAC30679.1; GSPDB:GN00147
C;Genetics:
A;Gene: ML1726
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 3.3%; Score 7; DB 2; Length 305;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 LSDVLTA 188
Db 264 LSDVLTA 270

RESULT 33
T46938
ornithine cyclodeaminase (EC 4.3.1.12) [similarity] - Agrobacterium tumefaciens plasmid I
C;Species: Agrobacterium tumefaciens
C;Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 02-Sep-2000
C;Accession: T46938
R;Oger, P.-M.; Farraud, S.K.
submitted to the EMBL Data Library, May 1998
A;Reference number: Z24140
A;Accession: T46938
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-307 <OGE>
A;Cross-references: EMBL:AF065244; PIDN:AAD31597.1
A;Experimental source: strain K84
C;Genetics:
A;Gene: ocd
A;Genome: plasmid pATK84b
C;Superfamily: ornithine cyclodeaminase
C;Keywords: ammonia-lyase; carbon-nitrogen lyase

Query Match 3.3%; Score 7; DB 2; Length 307;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 FGVLSDV 185
Db 50 FGVLSDV 56

RESULT 34
AD0243
hypothetical protein YPO1995 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AD0243
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, I
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AD0243
A;Status: preliminary
```

A:Molecule type: DNA
A:Residues: 1-310 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC90808.1; PID:gl5980009; GSPDB:GN00175
C:Genetics:
A:Gene: YP01995

Query Match 3.3%; Score 7; DB 2; Length 310;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 181 VLSDVLT 187
| | | | |
Db 278 VLSDVLT 284

RESULT 35
A72371
ClpOpeptide ABC transporter, ATP-binding protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: A72371
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: A72371
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333 <ARN>
A:Cross-references: GB:AE001726; GB:AE000512; NID:g4980992; PIDN:AAD35583.1; PID:g498100
C:Genetics:
A:Experimental source: strain MSB8
A:Gene: TMO498
C:Superfamily: inner membrane protein malk; ATP-binding cassette homology
P;27-234/Domain: ATP-binding cassette homology <ABC>

Query Match 3.3%; Score 7; DB 2; Length 333;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 95 EEGIVKA 101
| | | | |
Db 20 EEGIVKA 26

RESULT 36
G97024
probable phosphoesterase (EC 3.1.1.-) CAC1010 [similarity] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 22-Oct-2001
C:Accession: G97024
R:Nolling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: G97024
A:Molecule type: DNA
A:Residues: 1-345 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK78986.1; PID:gl5023919; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Comment: This sequence has motifs characteristic of a variety of phosphoesterases.
C:Genetics:
A:Gene: CAC1010
C:Superfamily: Clostridium acetobutylicum probable phosphoesterase CAC1078; phosphoesterase
F;33-120/Domain: phosphoesterase core homology <PEC>

Query Match 3.3%; Score 7; DB 1; Length 345;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LLIAAVA 11
| | | | |
Db 10 LLIAAVA 16

RESULT 37

Tl1130
hypothetical protein T30A10.70 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: Tl1130
R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Bar
submitted to the Protein Sequence Database, September 1999

A:Reference number: Z18708

A:Accession: Tl1130

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-349 <BEV>

A:Cross-references: EMBL:AL117386

A:Experimental source: cultivar Columbia; BAC clone T30A10

C:Genetics:

A:Map position: 4

A:Introns: 87/2; 140/2; 185/3; 211/2; 258/3; 281/3

A>Note: T30A10.70

Query Match 3.3%; Score 7; DB 2; Length 349;

Best Local Similarity 100.0%; Pred. No. 65;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 IDDAIAA 39

| | | | |

Db 265 IDDAIAA 271

RESULT 38

DIAG

ornithine cyclodeaminase (EC 4.3.1.12) [validated] - Agrobacterium sp. plasmid pTIC58
C:Species: Agrobacterium sp.
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 17-Mar-2000

C:Accession: S00402

R:Sans, N.; Schindler, U.; Schroeder, J.

Eur. J. Biochem. 173, 123-130, 1988

A:Title: Ornithine cyclodeaminase from Ti plasmid c58. DNA sequence, enzyme properties ar

A:Reference number: S00402; MUID:88185308; PMID:3281832

A:Accession: S00402

A:Molecule type: DNA

A:Residues: 1-354 <SAN>

A:Cross-references: EMBL:X07435; NID:g99107; PIDN:CAA30316.1; PID:g99108

C:Genetics:

A:Gene: ocd

A:Genome: plasmid pTIC58

C:Function:

A:Description: EC 4.3.1.12 [validated, MUID:88185308]; ornithine cyclodeaminase; catalyz

A>Note: regulated by L-arginine

A>Note: cofactor NAD+

C:Superfamily: ornithine cyclodeaminase

C:Keywords: ammonia-lyase; carbon-nitrogen lyase; opine catabolism

Query Match 3.3%; Score 7; DB 1; Length 354;

Best Local Similarity 100.0%; Pred. No. 66;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 179 FGVLSDV 185

| | | | |

Db 96 FGVLSDV 102

RESULT 39

S55582

ornithine cyclodeaminase (EC 4.3.1.12) - Agrobacterium tumefaciens

C:Species: Agrobacterium tumefaciens

C:Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 21-Jul-2000

C;Accession: S55582
R;Zanker, H.; Lurz, G.; Langridge, U.; Langridge, P.; Schroeder, J.
submitted to the EMBL Data Library, February 1994
A;Description: Octopine and nopaline oxidases from Ti plasmids of Agrobacterium tumefaciens
A;Reference number: S55578
A;Accession: S55582
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-354 <ZAN>
A;Cross-references: EMBL:Z30316; NID:g496533; PIDN:CAA82966.1; PID:g496538
C;Superfamily: ornithine cyclodeaminase
C;Keywords: ammonia-lyase; carbon-nitrogen lyase

Query Match 3.3%; Score 7; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 FGVLSDV 185
|||||
Db 96 FGVLSDV 102

RESULT 40
A32049
ornithine cyclodeaminase (EC 4.3.1.12) - Agrobacterium tumefaciens plasmid pTiAch5
C;Species: Agrobacterium tumefaciens
C;Date: 13-Jul-1989 #sequence_revision 13-Jul-1989 #text_change 16-Jul-1999
C;Accession: A32049; S55590
R;Schindler, U.; Sans, N.; Schroeder, J.
J.; Bacteriol. 171, 847-854, 1989
A;Title: Ornithine cyclodeaminase from octopine Ti plasmid Ach5: identification, DNA sequence
A;Reference number: A32049; MUID:89123162; PMID:2644238
A;Accession: A32049
A;Molecule type: DNA
A;Residues: 1-356 <SCH>
A;Cross-references: GB:M24146; NID:g154777; PIDN:AAA50518.1; PID:g154778
R;Zanker, H.; Lurz, G.; Langridge, U.; Langridge, P.; Schroeder, J.
submitted to the EMBL Data Library, February 1994
A;Description: Octopine and nopaline oxidases from Ti plasmids of Agrobacterium tumefaciens
A;Reference number: S55578
A;Accession: S55590
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-356 <ZAN>
A;Cross-references: EMBL:Z30328; NID:g496539; PIDN:CAA82989.1; PID:g496547
C;Genetics:
C;Genome: plasmid
C;Superfamily: ornithine cyclodeaminase
C;Keywords: ammonia-lyase; carbon-nitrogen lyase

Query Match 3.3%; Score 7; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 FGVLSDV 185
|||||
Db 96 FGVLSDV 102

RESULT 41
B95999
probable iron ABC transporter permease protein SMB21430 [imported] - Sinorhizobium meliloti
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: B95999
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: B95999
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-356 <KUR>

A;Cross-references: GB:AL591985; PIDN:CA49658.1; PID:g15141145; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMB21430
A;Genome: plasmid

Query Match 3.3%; Score 7; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 LSDVLTA 188
|||||
Db 53 LSDVLTA 59

RESULT 42
AF3229
ornithine cyclodeaminase [imported] - Agrobacterium tumefaciens (strain C58, Dupont) plac
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AF3229
R;Wood, D.W.; Stetbal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erge, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AF3229
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-371 <KUR>
A;Cross-references: GB:AE008690; PIDN:AAL46252.1; PID:g17744031; GSPDB:GN00189
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: ocd
A;Genome: plasmid
C;Superfamily: ornithine cyclodeaminase

Query Match 3.3%; Score 7; DB 2; Length 371;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 FGVLSDV 185
|||||
Db 113 FGVLSDV 119

RESULT 43
T15356
hypothetical protein B0563.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T15356
R;Favell, T.
submitted to the EMBL Data Library, June 1995
A;Description: The sequence of C. elegans cosmid B0563.
A;Reference number: Z16336
A;Accession: T15356
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-375 <FAV>
A;Cross-references: EMBL:U28740; NID:g861281; PID:g861282; PIDN:AAA68317.1; CESP:B0563.6

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:B0563.6

A:Introns: 140/3; 190/3; 221/3; 311/3

Query Match 3.3%; Score 7; DB 2; Length 375;

Best Local Similarity 100.0%; Pred. No. 69;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 139 VVALSLE 145

Db 62 VVALSLE 68

RESULT 44

E84743

hypothetical protein At2g33280 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: E84743

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

guss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: E84743

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-408 <STO>

A:Cross-references: GB:AE002093; NID:g2459411; PIDN:AAB80646.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g33280

A:Map position: 2

Query Match

Best Local Similarity 3.3%; Score 7; DB 2; Length 408;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 180 GVLSDDL 186

Db 8 GVLSDDL 14

RESULT 45

D96664

hypothetical protein T12P18.5 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: D96664

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Mafti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: D96664

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-415 <STO>

A:Cross-references: GB:AE005173; NID:g6358783; PIDN:AAF07364.1; GSPDB:GN00141

C:Genetics:

A:Gene: T12P18.5

A:Map position: 1

Query Match

Best Local Similarity 3.3%; Score 7; DB 2; Length 415;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 139 VVALSLE 145

Db 173 VVALSLE 179

RESULT 46

T17123

hypothetical protein T30A10.10 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T17123

R:Bevan, M.; Van der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Bar

submitted to the Protein Sequence Database, September 1999

A:Reference number: Z18708

A:Accession: T17123

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-427 <BEV>

A:Cross-references: EMBL:AL117386

A:Experimental source: cultivar Columbia; BAC clone T30A10

C:Genetics:

A:Map position: 4

A:Introns: 57/3; 165/2; 218/2; 263/3; 289/2; 336/3; 359/3

A:Note: T30A10.10

Query Match

Best Local Similarity 3.3%; Score 7; DB 2; Length 427;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 IDDAIAA 39

Db 343 IDDAIAA 349

RESULT 47

C97237

histidyl-tRNA synthetase [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C:Accession: C97237

R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A:Reference number: A56900; MUID:21359325; PMID:21359325

A:Accession: C97237

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-430 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK80686.1; PID:g15025776; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC2740

Query Match

Best Local Similarity 3.3%; Score 7; DB 2; Length 430;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 YDKITEE 28

Db 192 YDKITEE 198

RESULT 48

C86178

hypothetical protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: C86178

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
Xer, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: C86178
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-431 <STO>
A:Cross-references: GB:AE005172; NID:92494120; PIDN:AB80629.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 3.3%; Score 7; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 180 GVLSVVL 186
| | | | |
Db 97 GVLSVVL 103

RESULT 49
xanthine permease [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 17-May-2002
C:Accession: C86768
R:Botolin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: C86768
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-434 <STO>
A:Cross-references: GB:AE005176; PID:G12724110; PIDN:AAK05245.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: pbux
C:Superfamily: hypothetical protein b2882

Query Match 3.3%; Score 7; DB 2; Length 434;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 LLIAAVA 11
| | | | |
Db 372 LLIAAVA 378

RESULT 50
B53764
beta2-chimerin, cerebellar - rat (fragment)
N:Alternate names: GTPase-activating protein
C:Species: Rattus norvegicus (Norway rat)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Mar-1999
C:Accession: B53764
R:Leung, T.; How, B.E.; Manser, E.; Lim, L.
J. Biol. Chem. 269, 12888-12892, 1994
A:Title: Cerebellar beta2-chimaerin, a GTPase-activating protein for p21 Ras-related Rac
A:Reference number: A53764; MUID:94230370; PMID:8175705
A:Accession: B53764
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-443 <LEU>
A:Superfamily: protein kinase C zinc-binding repeat homology; SH2 homology
F:34-114/Domain: SH2 homology <SH2>
F:190-239/Domain: protein kinase C zinc-binding repeat homology <K22>

Query Match 3.3%; Score 7; DB 2; Length 446;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 130 HVISDIQ 136
| | | | |
Db 440 HVISDIQ 446

RESULT 51
A53764
beta2-chimerin, cerebellar - human
N:Alternate names: GTPase-activating protein
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
C:Accession: A53764
R:Leung, T.; How, B.E.; Manser, E.; Lim, L.
J. Biol. Chem. 269, 12888-12892, 1994
A:Title: Cerebellar beta2-chimaerin, a GTPase-activating protein for p21 Ras-related Rac
A:Reference number: A53764; MUID:94230370; PMID:8175705
A:Accession: A53764
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-466 <LEU>
A:Cross-references: GB:L29126; NID:9457229; PIDN:AAA19191.1; PID:9457230
C:Superfamily: protein kinase C zinc-binding repeat homology; SH2 homology
F:57-137/Domain: SH2 homology <SH2>
F:213-262/Domain: protein kinase C zinc-binding repeat homology <K22>

Query Match 3.3%; Score 7; DB 2; Length 466;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 76 IEARGLK 82
| | | | |
Db 299 IEARGLK 305

RESULT 52
T01347
pectinesterase homolog F6N15.23 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 26-Aug-1999
C:Accession: T01347
R:Ryan, E.; Edwards, J.; Pape, K.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of A. thaliana F6N15.
A:Reference number: Z14297
A:Accession: T01347
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-474 <RYA>
A:Cross-references: EMBL:AF069299; NID:G3193311; PID:G3193313
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4
A:Introns: 138/3; 248/1
A:Note: F6N15.23
C:Superfamily: pectinesterase

Query Match 3.3%; Score 7; DB 2; Length 443;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 76 IEARGLK 82
| | | | |
Db 276 IEARGLK 282

RESULT 51
A53764
beta2-chimerin, cerebellar - human
N:Alternate names: GTPase-activating protein
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
C:Accession: A53764
R:Leung, T.; How, B.E.; Manser, E.; Lim, L.
J. Biol. Chem. 269, 12888-12892, 1994
A:Title: Cerebellar beta2-chimaerin, a GTPase-activating protein for p21 Ras-related Rac
A:Reference number: A53764; MUID:94230370; PMID:8175705
A:Accession: A53764
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-466 <LEU>
A:Cross-references: GB:L29126; NID:9457229; PIDN:AAA19191.1; PID:9457230
C:Superfamily: protein kinase C zinc-binding repeat homology; SH2 homology
F:57-137/Domain: SH2 homology <SH2>
F:213-262/Domain: protein kinase C zinc-binding repeat homology <K22>

Query Match 3.3%; Score 7; DB 2; Length 466;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 76 IEARGLK 82
| | | | |
Db 299 IEARGLK 305

RESULT 52
T01347
pectinesterase homolog F6N15.23 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 26-Aug-1999
C:Accession: T01347
R:Ryan, E.; Edwards, J.; Pape, K.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of A. thaliana F6N15.
A:Reference number: Z14297
A:Accession: T01347
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-474 <RYA>
A:Cross-references: EMBL:AF069299; NID:G3193311; PID:G3193313
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4
A:Introns: 138/3; 248/1
A:Note: F6N15.23
C:Superfamily: pectinesterase

Query Match 3.3%; Score 7; DB 2; Length 474;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 130 HVISDIQ 136
| | | | |
Db 440 HVISDIQ 446

RESULT 53
AE0274
L-arabinose isomerase (EC 5.3.1.4) [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis

C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
 C;Accession: AE0274
 E;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittball, R.W.; Holden, M.T.G.; Prentice, M.B.; Genot-farraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Li, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
 A;Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A;Reference number: AB0001; MUID:21470413; PMID:11586360
 A;Accession: AE0274
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-500 <KUR>
 A;Cross-references: GB:AL590842; PIDN:CAC91057.1; PID:g15980249; GSPDB:GN00175
 C;Genetics:
 A;Gene: araA
 C;Superfamily: L-arabinose isomerase
 C;Keywords: intramolecular oxidoreductase; isomerase

Query Match 3.3%; Score 7; DB 2; Length 500;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 IGGLSIL 175
 DB 90 IGGLSIL 96

RESULT 54
 T40077
 hypothetical protein SPBC29A3.06 - fission yeast (*Schizosaccharomyces pombe*)
 C;Species: *Schizosaccharomyces pombe*
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C;Accession: T40077
 R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.
 A;Reference number: T21904
 A;Accession: T40077
 A;Status: Preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-556 <LYN>
 A;Cross-references: EMBL:AL022299; PIDN:CAAL18383.1; GSPDB:GN000667; SPDB:SPBC29A3.06
 A;Experimental source: strain 972h-; cosmid c29A3
 C;Genetics:
 A;Gene: SPDB:SPBC29A3.06
 A;Map position: 2

Query Match 3.3%; Score 7; DB 2; Length 556;
 Best Local Similarity 100.0%; Pred. No. 97;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GVLSDVL 186
 DB 350 GVLSDVL 356

RESULT 55
 H85731
 Rhs element associated protein Z2259 [imported] - *Escherichia coli* (strain O157:H7, subs
 C;Species: *Escherichia coli*
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C;Accession: H85731
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A;Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A;Reference number: H85480; MUID:21074935; PMID:11206551
 A;Accession: H85731
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-572 <STO>
 A;Cross-references: GB:AE005174; NID:g12515237; PIDN:AA056316.1; GSPDB:GN00145; UWGP:Z22
 A;Experimental source: strain O157:H7, substrain EDL933
 C;Genetics:

A;Gene: Z2259

Query Match 3.3%; Score 7; DB 2; Length 572;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 SNEYDLA 119
 DB 499 SNEYDLA 505

RESULT 56

AE3035

conserved hypothetical protein Atu3891 [imported] - *Agrobacterium tumefaciens* (strain C58)
 C;Species: *Agrobacterium tumefaciens*
 C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C;Accession: AE3035
 R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan,
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A;Title: Genome sequence of *Agrobacterium tumefaciens* strain C58.
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.
 ster, E.W.
 A;Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
 A;Reference number: AB2577; MUID:21608550; PMID:11743193
 A;Accession: AE3035
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-573 <KUR>
 A;Cross-references: GB:AB008689; PIDN:AAL44699.1; PID:g17742328; GSPDB:GN00187
 A;Experimental source: strain C58 (Dupont)
 C;Genetics:
 A;Gene: Atu3891
 A;Map position: linear chromosome

Query Match 3.3%; Score 7; DB 2; Length 573;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 KVLAPAF 206
 DB 80 KVLAPAF 86

RESULT 57

F98250

hypothetical protein AGR_L1912 [imported] - *Agrobacterium tumefaciens* (strain C58, Cerc
 C;Species: *Agrobacterium tumefaciens*
 C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
 C;Accession: F98250
 R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A;Title: Genome sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*
 A;Reference number: A97359; MUID:21608551; PMID:11743194
 A;Accession: F98250
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-573 <KUR>
 A;Cross-references: GB:AB007870; PIDN:AAK89528.1; PID:g15159407; GSPDB:GN00170
 C;Genetics:
 A;Gene: AGR_L1912
 A;Map position: linear chromosome

Query Match 3.3%; Score 7; DB 2; Length 573;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 KVLAPAF 206
 DB 80 KVLAPAF 86

```
RESULT 58
A:Molecule type: DNA
A:Residues: 1-657 <COL>
A:Cross-references: GB:Z96072; GB:AL123456; NID:g3261793; PIDN:CAB09497.1; PID:ei300032;
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV2690c

Query Match      3.3%; Score 7; DB 2; Length 657;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4  FLIIAAV 10
        |||||
Db      324 FLIIAAV 330

RESULT 61
H75460
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C:Accession: H75460
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
  M.; Shen, H.O.; Venter, J.C.; Fraser, C.M.
  S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
  Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: H75460
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-736 <WHI>
A:Cross-references: GB:AE001943; GB:AE000513; NID:g6458624; PIDN:AAFI0482.1; PID:g6458624
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0905
A:Map position: 1
C:Superfamily: Haloferax hypothetical protein 4 (gyrB region)

Query Match      3.3%; Score 7; DB 2; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      100 KAHLIG 106
        |||||
Db      666 KAHLIG 672

RESULT 62
S12858
vira protein - Agrobacterium rhizogenes plasmid pRiA4
C:Species: Agrobacterium rhizogenes
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 17-Nov-2000
C:Accession: S12858; S08133
R:Endoh, H.; Hirayama, T.; Aoyama, T.; Oka, A.
  FEBS Lett. 271, 28-32, 1990
A:Title: Characterization of the vira gene of the agropine-type plasmid pRiA4 of Agrobact
A:Reference number: S12858; MUID:91032080; PMID:2226811
A:Accession: S12858
A:Molecule type: DNA
A:Residues: 1-829 <END>
A:Cross-references: EMBL:X51418; NID:g38991; PIDN:CAA35780.1; PID:g38992
R:Aoyama, T.; Takanami, M.; Oka, A.
  Nucleic Acids Res. 17, 8711-8725, 1989
A:Title: Signal structure for transcriptional activation in the upstream regions of viru
A:Reference number: S08133; MUID:90067840; PMID:2479910
A:Accession: S08133
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-16 <AOY>
A:Cross-references: EMBL:X15909; NID:g38977; PIDN:CAA34025.1; PID:g38978
C:Genetics:
A:Gene: vira
```


A;Genome: plasmid pR1A4b

Query Match 3.3%; Score 7; DB 2; Length 829;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 141 ALSLEIS 147

Db 181 ALSLEIS 187

RESULT 53

S06972

vira protein - Agrobacterium tumefaciens plasmid pTIC58

C;Species: Agrobacterium tumefaciens

C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 17-Nov-2000

C;Accession: S06972; S11825

R;Morel, P.; Powell, B.S.; Rogowsky, P.M.; Kado, C.I.

Mol. Microbiol. 3: 1237-1246, 1989

A;Title: Characterization of the vira virulence gene of the nopaline plasmid, pTIC58, of

A;Reference number: S06972; MUID: 90014184; PMID: 2796735

A;Accession: S06972

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-833 <MOR>

R;Rogowsky, P.M.; Powell, B.S.; Shirasu, K.; Lin, T.S.; Morel, P.; Zyprian, E.M.; Steck,

Plasmid 23: 85-106, 1990

A;Title: Molecular characterization of the vir regulon of Agrobacterium tumefaciens: com

A;Reference number: S11825; MUID: 90301800; PMID: 2194232

A;Accession: S11825

A;Status: preliminary; translation not shown

A;Molecule type: DNA

A;Residues: 1-550, 'L', 552-654, 'SC', 657-833 <ROG>

A;Cross-references: EMBL:J03320; NID: g154781; PIDN: AAA91590.1; PID: g154782

C;Genetics:

A;Gene: vira

A;Genome: plasmid pTIC58

Query Match

Best Local Similarity 3.3%; Score 7; DB 2; Length 833;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 141 ALSLEIS 147

Db 181 ALSLEIS 187

RESULT 54

AD3248

two component sensor kinase vira [imported] - Agrobacterium tumefaciens (strain C58, Dup

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C;Accession: AD3248

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; MUID: 21608550; PMID: 11743193

A;Accession: AD3248

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-833 <KUR>

A;Cross-references: GB:AE008690; PIDN: AAL46402.1; PID: g17744196; GSPDB: GN00189

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: vira

A;Genome: plasmid

Query Match

Best Local Similarity 3.3%; Score 7; DB 2; Length 833;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 141 ALSLEIS 147

Db 181 ALSLEIS 187

RESULT 65

H72597

hypothetical protein APE1248 - Aeropyrum pernix (strain KI)

C;Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000

C;Accession: H72597

R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; K

DNA Res. 6: 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A;Reference number: A72450; MUID: 99310339; PMID: 10382966

A;Accession: H72597

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-871 <KAW>

A;Cross-references: DDBJ:AP000061; NID: g5104821; PIDN: BAA80239.1; PID: d1044024; PID: g5104

A;Experimental source: strain KI

C;Genetics:

A;Gene: APE1248

C;Superfamily: Aeropyrum pernix hypothetical protein APE1248

Query Match 3.3%; Score 7; DB 2; Length 871;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 169 IGGLSIL 175

Db 802 IGGLSIL 808

RESULT 66

C75498

probable metalloproteinase - Deinococcus radiodurans (strain R1)

C;Species: Deinococcus radiodurans

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C;Accession: C75498

R;White, O.; Eisen, J.A.; Heidelberg, J.P.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; F

; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A;Reference number: A75250; MUID: 20036896; PMID: 10567266

A;Accession: C75498

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-996 <WHI>

A;Cross-references: GB:AE001919; GB:AE00513; NID: g6458307; PIDN: AAF10194.1; PID: g6458314

A;Experimental source: strain R1

C;Genetics:

A;Gene: DR0617

A;Map position: 1

Query Match 3.3%; Score 7; DB 2; Length 996;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 180 GVLSDVL 186

Db 328 GVLSDVL 334

RESULT 67

H90861

probable multidrug-efflux transport protein ECs1864 [imported] - Escherichia coli (strain

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C:Accession: H90861
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: H90861
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1027 <HAY>
A:Cross-references: GS:BA000007; PIDN:BA03361329; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 050952
C:Genetics:
A:Gene: ECs1864
C:Superfamily: acriflavin resistance protein

Query Match 3.3%; Score 7; DB 2; Length 1027;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 FLIIAAV 10
|||||
Db 135 FLIIAAV 141

RESULT 68
D85757
probable efflux pump Z2508 [imported] - *Escherichia coli* (strain O157:H7, substrain EDL9
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: D85757
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85757
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1027 <STO>
A:Cross-references: GB:AE005174; NID:G12515493; PIDN:AA056520.1; GSPDB:GN00145; UWGP:Z25
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z2508
C:Superfamily: acriflavin resistance protein

Query Match 3.3%; Score 7; DB 2; Length 1027;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 FLIIAAV 10
|||||
Db 135 FLIIAAV 141

RESULT 69
B71720
hypothetical protein RP108 - *Rickettsia prowazekii*
C:Species: *Rickettsia prowazekii*
C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
R:Anderson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichert-Ponten, T.; Alsmark, U
Nature 396, 133-140, 1998
A:Title: The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499; PMID:9823893
A:Accession: B71720
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1155 <AND>
A:Cross-references: GB:AJ235270; GB:AJ235269; NID:G3860572; PIDN:CAA14577.1; PID:el34242
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: RP108

Query Match 3.3%; Score 7; DB 2; Length 1155;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 KITEIN 30
|||||
Db 1078 KITEIN 1084

RESULT 70
G70837
probable ABC transporter Rv0194 - *Mycobacterium tuberculosis* (strain H37RV)
C:Species: *Mycobacterium tuberculosis*
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: G70837
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: G70837
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1194 <COL>
A:Cross-references: GB:AL021928; GB:AL123456; NID:G261522; PIDN:CAA17316.1; PID:G290959;
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv0194
C:Superfamily: *Mycobacterium tuberculosis* probable ABC transporter Rv0194; ATP-binding ca
C:Keywords: ATP
F:350-544/Domain: ATP-binding cassette homology <ABC>

Query Match 3.3%; Score 7; DB 1; Length 1194;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 IDDAIAA 39
|||||
Db 45 IDDAIAA 51

RESULT 71
H90698
Rsd core protein with extension [imported] - *Escherichia coli* (strain O157:H7, substrain
C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: H90698
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genom
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: H90698
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1398 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA03361329.1; PID:G13360018; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 050952
C:Genetics:
A:Gene: ECs0560
C:Superfamily: rtsP protein

Query Match 3.3%; Score 7; DB 2; Length 1398;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 SMEYDLA 119
|||||
Db 641 SMEYDLA 647

RESULT 72

B85549
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: B85549
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11204551
 A:Accession: B85549
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1398 <STO>
 A:Cross-references: GB:AE005174; NID:gl2513390; PIDN:AA054854.1; GSPDB:GN00145; UWGP:206
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 C:Superfamily: rhsF protein

Query Match 3.3%; Score 7; DB 2; Length 1398;

Best Local Similarity 100.0%; Pred. No. 2.1e+02; Mismatches 0; Indels 0; Gaps 0;

QY 113 SMEYDLA 119
 |||||
 Db 641 SMEYDLA 647

RESULT 73

E90886
 RhsE core protein with extension [imported] - Escherichia coli (strain O157:H7, substrain
 C:Species: Escherichia coli
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: E90886
 R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: E90886
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1400 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA035484.1; PID:gl3361527; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 050952
 C:Genetics:
 A:Gene: Ecs2061
 C:Superfamily: rhsF protein

Query Match 3.3%; Score 7; DB 2; Length 1400;

Best Local Similarity 100.0%; Pred. No. 2.1e+02; Mismatches 0; Indels 0; Gaps 0;

QY 113 SMEYDLA 119
 |||||
 Db 641 SMEYDLA 647

RESULT 74

S49764
 SEC7 protein - yeast (Saccharomyces cerevisiae)
 N;Alternate names: protein YD9489.05c; protein YDR170C
 C:Species: Saccharomyces cerevisiae
 C>Date: 13-Jan-1995 #sequence_revision 23-Aug-1996 #text_change 21-Jul-2000
 C:Accession: S49764; S50916; A31068; A28784
 R;Murphy, L.; Harris, D.E.
 submitted to the EMBL Data Library, November 1994
 A:Reference number: S49764
 A:Accession: S49764
 A:Molecule type: DNA

A:Residues: 1-542 <MUR>
 A:Cross-references: EMBL:Z46727; NID:gl289283; PID:e223636; PID:gl204159; MIPS:YDR170C
 R;Oliver, K.; Harris, D.
 submitted to the EMBL Data Library, January 1994
 A:Reference number: S50912
 A:Accession: S50916
 A:Molecule type: DNA
 A:Residues: 204-2009 <OLT>
 A:Cross-references: EMBL:Z47813; NID:ig642274; PID:e135579; PID:gl326010; MIPS:YDR170C
 R;Achetter, T.; Franzusoff, A.; Field, C.; Schekman, R.
 J. Biol. Chem. 263, 11711-11717, 1988
 A>Title: SEC7 encodes an unusual, high molecular weight protein required for membrane tra
 A:Reference number: A31068; MUID:89298841; PMID:3042778
 A:Accession: A31068
 A:Molecule type: DNA
 A:Residues: 1-187,'S',189-398,'LL',399,'C',403-1030,'PAICF',1031,'L',1038-2009 <ACH1>
 A:Cross-references: EMBL:J03918; NID:gl435186; PIDN:AA04031.1; PID:gl172570
 R;Achetter, T.; Franzusoff, A.; Field, C.; Schekman, R.
 submitted to the Protein Sequence Database, August 1988
 A:Reference number: A94619
 A:Accession: A28784
 A:Molecule type: DNA
 A:Residues: 1-187,'S',189-398,'LL',399,'C',403-1030,'PAICF',1031,'L',1038-2009 <ACH2>
 C:Genetics:
 A:Gene: SGD:SEC7
 A:Cross-references: SGD:S0002577; MIPS:YDR170C
 A:Map position: 4R
 C:Keywords: transmembrane protein
 F:582-598/Domain: transmembrane #status predicted <TM1>
 F:728-744/Domain: transmembrane #status predicted <TM2>
 F:1817-1833/Domain: transmembrane #status predicted <TM3>

Query Match 3.3%; Score 7; DB 2; Length 2009;

Best Local Similarity 100.0%; Pred. No. 2.9e+02; Mismatches 0; Indels 0; Gaps 0;

QY 177 PIFGVLS 183
 |||||
 Db 1588 PIFGVLS 1594

RESULT 75

T30849
 actin binding protein ACF7, neural isoform 3 - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
 C:Accession: T30849
 R;Bernier, G.; Mathieu, M.; De Repentigny, Y.; Vidal, S.M.; Kothary, R.
 Genomics 38, 19-29, 1996
 A>Title: Cloning and characterization of mouse ACF7, a novel member of the dystonin subf
 A:Reference number: Z20900; MUID:97124842; PMID:8954775
 A:Accession: T30849
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-2033 <BER>
 A:Cross-references: EMBL:U67205; NID:gl675225; PID:gl675226; PIDN:AA052990.1
 C:Genetics:
 A:Gene: mACF7
 A:Map position: 4

Query Match

Best Local Similarity 100.0%; Pred. No. 2.9e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 LSLRISD 148
 |||||
 Db 329 LSLRISD 335

RESULT 76

S71457
 NADH2 dehydrogenase (ubiquinone) (BC 1.6.5.3) chain 6 - rice mitochondrion (fragment)
 C:Species: mitochondrion Oryza sativa (rice)

C>Date: 29-Jan-1998 #sequence_revision 13-Feb-1998 #text_change 03-Jun-2002
C/Accession: S71457
R/Nakazono, M.; Ito, Y.; Tsutsumi, N.; Hirai, A.
Curr. Genet. 29, 412-416, 1996
A/Title: The gene for a subunit of an ABC-type heme transporter is transcribed together
A/Reference number: S71456; MUID:96207463; PMID:8625418
A/Accession: S71457
A/Molecule type: DNA
A/Residues: 1-34 <NAK>
A/Cross-references: ENBL:D64067; NID:gl395187; PID:BAA10943.1; PID:gl395189
A/Experimental source: strain Nipponbare; tissue type leaf
C/Genetics:
A/Gene: nad6
A/Genome: mitochondrion
A/Superfamily: NADH dehydrogenase (ubiquinone) chain 6
C/Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;
Query Match 2.8%; Score 6; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 202 LAPAFK 207
DB 4 LAPAFK 9
RESULT 77
G84147
Hypothetical protein BH3983 [imported] - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C/Accession: G84147
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83550; MUID:20512582; PMID:11058132
A/Accession: G84147
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-34 <STO>
A/Cross-references: GB:AP001520; GB:BA000004; NID:gl0176401; PIDN:BA07702.1; GSPDB:GN00
A/Experimental source: strain C-125
C/Genetics:
A/Gene: BH3983
Query Match 2.8%; Score 6; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 207 KRELEK 212
DB 27 KRELEK 32
RESULT 78
S23286
light-harvesting protein beta chain - Ectothiorhodospira halochloris (fragment)
N/Alternate names: antenna pigment protein beta chain
C/Species: Ectothiorhodospira halochloris
C/Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 21-Aug-1998
C/Accession: S23286
R/Wagner-Huber, R.; Brunisholz, R.A.; Bissig, I.; Frank, G.; Suter, F.; Zuber, H.
Eur. J. Biochem. 205, 917-925, 1992
A/Title: The primary structure of the antenna polypeptides of Ectothiorhodospira halochl
A/Reference number: S23164; MUID:92249336; PMID:1577009
A/Accession: S23286
A/Molecule type: protein
A/Residues: 1-40 <WAG>
C/Superfamily: light-harvesting protein beta chain
C/Keywords: antenna complex; bacteriochlorophyll; light-harvesting polypeptide; magnesium
Query Match 2.8%; Score 6; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 1e+02;
C/Date: 29-Jan-1998 #sequence_revision 13-Feb-1998 #text_change 03-Jun-2002
C/Accession: S71457
R/Nakazono, M.; Ito, Y.; Tsutsumi, N.; Hirai, A.
Curr. Genet. 29, 412-416, 1996
A/Title: The gene for a subunit of an ABC-type heme transporter is transcribed together
A/Reference number: S71456; MUID:96207463; PMID:8625418
A/Accession: S71457
A/Molecule type: DNA
A/Residues: 1-34 <NAK>
A/Cross-references: ENBL:D64067; NID:gl395187; PID:BAA10943.1; PID:gl395189
A/Experimental source: strain Nipponbare; tissue type leaf
C/Genetics:
A/Gene: nad6
A/Genome: mitochondrion
A/Superfamily: NADH dehydrogenase (ubiquinone) chain 6
C/Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;
Query Match 2.8%; Score 6; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 202 LAPAFK 207
DB 4 LAPAFK 9
RESULT 77
G84147
Hypothetical protein BH3983 [imported] - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C/Accession: G84147
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83550; MUID:20512582; PMID:11058132
A/Accession: G84147
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-34 <STO>
A/Cross-references: GB:AP001520; GB:BA000004; NID:gl0176401; PIDN:BA07702.1; GSPDB:GN00
A/Experimental source: strain C-125
C/Genetics:
A/Gene: BH3983
Query Match 2.8%; Score 6; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 207 KRELEK 212
DB 27 KRELEK 32
RESULT 78
S23286
light-harvesting protein beta chain - Ectothiorhodospira halochloris (fragment)
N/Alternate names: antenna pigment protein beta chain
C/Species: Ectothiorhodospira halochloris
C/Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 21-Aug-1998
C/Accession: S23286
R/Wagner-Huber, R.; Brunisholz, R.A.; Bissig, I.; Frank, G.; Suter, F.; Zuber, H.
Eur. J. Biochem. 205, 917-925, 1992
A/Title: The primary structure of the antenna polypeptides of Ectothiorhodospira halochl
A/Reference number: S23164; MUID:92249336; PMID:1577009
A/Accession: S23286
A/Molecule type: protein
A/Residues: 1-40 <WAG>
C/Superfamily: light-harvesting protein beta chain
C/Keywords: antenna complex; bacteriochlorophyll; light-harvesting polypeptide; magnesium
Query Match 2.8%; Score 6; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9 AVAFVA 14
DB 31 AVAFVA 36
RESULT 79
S68885
light-harvesting protein B885 beta-1 chain - Rhodocyclus tenuis
N/Alternate names: antenna/reaction-centre complex RC-B885 beta-1 chain
C/Species: Rhodocyclus tenuis
C/Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 30-Jan-1998
C/Accession: S68885
R/Hu, Q.; Brunisholz, R.A.; Frank, G.; Zuber, H.
Eur. J. Biochem. 238, 381-390, 1996
A/Title: The antenna complexes of the purple non-sulfur photosynthetic bacterium Rhodocy
A/Reference number: S68881; MUID:96283832; PMID:8681949
A/Accession: S68885
A/Molecule type: protein
A/Residues: 1-48 <HUQ>
A/Experimental source: DSM 109
C/Superfamily: light-harvesting protein beta chain
C/Keywords: antenna complex; bacteriochlorophyll; light-harvesting polypeptide; magnesium
Query Match 2.8%; Score 6; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 VAFVAV 15
DB 26 VAFVAV 31
RESULT 80
S68886
light-harvesting protein B885 beta-2 chain - Rhodocyclus tenuis
N/Alternate names: antenna/reaction-centre complex RC-B885 beta-2 chain
C/Species: Rhodocyclus tenuis
C/Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 30-Jan-1998
C/Accession: S68886
R/Hu, Q.; Brunisholz, R.A.; Frank, G.; Zuber, H.
Eur. J. Biochem. 238, 381-390, 1996
A/Title: The antenna complexes of the purple non-sulfur photosynthetic bacterium Rhodocy
A/Reference number: S68881; MUID:96283832; PMID:8681949
A/Accession: S68886
A/Molecule type: protein
A/Residues: 1-48 <HUQ>
A/Experimental source: DSM 109
C/Superfamily: light-harvesting protein beta chain
C/Keywords: antenna complex; bacteriochlorophyll; light-harvesting polypeptide; magnesium
Query Match 2.8%; Score 6; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 VAFVAV 15
DB 26 VAFVAV 31
RESULT 81
S55781
laminin alpha chain - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 04-Sep-1998
C/Accession: S55781
R/Kammerer, R.A.; Antonsson, P.; Schultness, T.; Fauser, C.; Engel, J.
J. Mol. Biol. 250, 64-73, 1995
A/Title: Selective chain recognition in the C-terminal alpha-helical coiled-coil region
A/Reference number: S55781; MUID:95326149; PMID:7602597
A/Accession: S55781
A/Status: Preliminary

A:Molecule type: protein
A:Residues: 1-49 <XAM>
C:Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-like h

Query Match 2.8%; Score 6; DB 2; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSAD 18
Db 30 VAVSAD 35

RESULT 82
A37354
sex-specific protein mst 316 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 19-Mar-1993 #sequence_revision 19-Mar-1993 #text_change 16-Feb-1997
C:Accession: A37354
R:DiBenedetto, A.J.; Harada, H.A.; Wolfner, M.F.
Dev. Biol. 139, 134-148, 1990
A:Title: Structure, cell-specific expression, and mating-induced regulation of a Drosoph
A:Reference number: A37354; MUID:90228604; PMID:2109712
A:Accession: A37354
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-52 <DIB>
A:Cross-references: GB:M32022
C:Genetics:
A:Gene: FlyBase:Acp95EF
A:Cross-references: FlyBase:FBgn0002863

Query Match 2.8%; Score 6; DB 2; Length 52;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVALSL 144
Db 13 VVALSL 18

RESULT 83
G90131
hypothetical protein orf62 [imported] - Guillardia theta nucleomorph
C:Species: nucleomorph Guillardia theta
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C:Accession: G90131
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Re
Nature 410, 1091-1095, 2001
A:Title: The highly reduced genome of an enslaved algal nucleus.
A:Reference number: A99082; MUID:11323671; PMID:11323671
A:Accession: G90131
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-62 <DOU>
A:Cross-references: GB:AF083031; NID:g13794369; PIDN:AAK39746.1; GSPDB:GN00152
C:Genetics:
A:Gene: orf62
A:Map position: 3
A:Genome: nucleomorph
C:Keywords: nucleomorph

Query Match 2.8%; Score 6; DB 2; Length 62;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 ALSLEI 146
Db 46 ALSLEI 51

RESULT 84

D84194
hypothetical protein Vng0352h [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: D84194
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: D84194
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-72 <STO>
A:Cross-references: GB:AF004437; NID:g10579974; PIDN:AAG18920.1; GSPDB:GN00138
C:Genetics:
A:Gene: VNG0352H

Query Match 2.8%; Score 6; DB 2; Length 72;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LIAAVA 11
Db 25 LIAAVA 30

RESULT 85
F75031
hydrogenase expression/formation protein (hycp) PAB7315 - Pyrococcus abyssi (strain Orsay
C:Species: Pyrococcus abyssi
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 14-Apr-2003
C:Accession: F75031
R:anonymous, Genoscope
Submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome struc
A:Reference number: A75001
A:Accession: F75031
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-75 <XAW>
A:Cross-references: GB:AJ248287; GB:AL096936; NID:g5458657; PIDN:CAB50147.1; PID:g5458661
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB7315
C:Superfamily: [NiFe]-hydrogenase maturation chaperone

Query Match 2.8%; Score 6; DB 2; Length 75;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 VGIVDF 67
Db 16 VGIVDF 21

RESULT 86
E69472
conserved hypothetical protein AF1782 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: E69472
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Attiach, P.; Kaine, B.P.; Sykes, S.N
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: E69472

A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Cross-references: GB:AE000979; GB:AE000782; NID:G2689302; PIDN:AAB89471.1; PID:G264876

Query Match 2.8%; Score 6; DB 2; Length 76;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 MNTEA 78
|||||
DB 30 MNTEA 35
|||||

RESULT 87
B33961
acyl-carrier protein acpA [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 28-Jul-2003
C;Accession: B33961
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: B33961
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-77 <STO>
A;Cross-references: GB:AP001515; GB:BA000004; NID:G10174886; PIDN:BA06209.1; GSPDB:GN00
A;Experimental source: strain C-125
C;Genetics:
A;Gene: acpA
C;Superfamily: acyl/peptidyl carrier protein; acyl carrier protein homology
C;Keywords: carrier protein

Query Match 2.8%; Score 6; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 LEISDE 149
|||||
DB 53 LEISDE 58
|||||

RESULT 88
H59420
Hydrogenase expression/formation protein (hypC) homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 14-Apr-2003
C;Accession: H59420
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
., Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea
A;Reference number: A59250; MUID:98049343; PMID:9389475
A;Accession: H59420
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-77 <KLE>
A;Cross-references: GB:AE001009; GB:AE000782; NID:G2689332; PIDN:AAB89878.1; PID:G264920
A;Superfamily: [NifH]-hydrogenase maturation chaperone

Query Match 2.8%; Score 6; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 IVDFKG 69
|||||
DB 18 IVDFKG 23
|||||

RESULT 89
E83102
conserved hypothetical protein PA4357 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: E83102
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bri
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Labig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: E83102
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-80 <STO>
A;Cross-references: GB:AE004851; GB:AE004091; NID:G9950571; PIDN:AAG07745.1; GSPDB:GN001
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA4357

Query Match 2.8%; Score 6; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 ARGLKQ 83
|||||
DB 16 ARGLKQ 21
|||||

RESULT 90
C82809
hypothetical protein XF0424 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: C82809
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: C82809
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-80 <SIM>
A;Cross-references: GB:AE003893; GB:AE003849; NID:G9105253; PIDN:AAF83334.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvares, R.; Al
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, D.M.; Carrez, H.
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohne
J.D.; Junqueira, M.B.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigt
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
A;Genetics:
A;Gene: XF0424

Query Match 2.8%; Score 6; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LIAAVA 11
|||||
DB 67 LIAAVA 72
|||||

RESULT 91

AF2125
hypothetical protein asl2557 [imported] - Nostoc sp. (strain PCC 7120)
C/Species: Nostoc sp. PCC 7120
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C/Accession: AF2125
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A/Reference number: AB1807; MUID:21595285; PMID:11759840
A/Accession: AF2125
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-80 <KUR>
A/Cross-references: GB:BA000019; PIDN:BA74256.1; PID:gl7131649; GSPDB:GN00179
A/Experimental source: strain PCC 7120
C/Genetics:
A/Gene: asl2557

Query Match 2.8%; Score 6; DB 2; Length 80;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 GREGIV 99

DB 38 GREGIV 43

RESULT 92

S73167
H+-transporting two-sector ATPase (EC 3.6.3.14) chain c - red alga (Porphyra purpurea)
C/Species: chloroplast Porphyra purpurea
C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 03-Jun-2002
C/Accession: S73167
R/Reith, M.; Munkholland, J.
Plant Mol. Biol. Rep. 13, 333-335, 1995
A/Title: Complete nucleotide sequence of the Porphyra purpurea chloroplast genome.
A/Reference number: S73108
A/Accession: S73167

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-82 <RET>

A/Cross-references: EMBL:U39804; NID:gl276652; PIDN:AC08132.1; PID:gl276712

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995

C/Genetics:

A/Gene: atpH

A/Genome: chloroplast

C/Superfamily: H+-transporting ATP synthase lipid-binding protein

C/Keywords: ATP biosynthesis; chloroplast; hydrolase; membrane-associated complex; thylakoid

Query Match

Best Local Similarity 2.8%; Score 6; DB 2; Length 82;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVALSL 144

DB 69 VVALSL 74

RESULT 93

S26958
H+-transporting two-sector ATPase (EC 3.6.3.14) lipid-binding protein - red alga (Antithamion sp.)
N/Alternate names: H+-transporting ATP synthase chain c; H+-transporting ATP synthase chain c
C/Species: chloroplast Antithamion sp.
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
C/Accession: S26958
R/Kostrzewa, M.; Zetsche, K.
J. Mol. Biol. 227, 961-970, 1992

A/Title: Large ATP synthase operon of the red alga Antithamion sp. resembles the corresponding

A/Reference number: S26957; MUID:93021132; PMID:1404401

A/Accession: S26958
A/Molecule type: DNA
A/Residues: 1-82 <KOS>
A/Cross-references: EMBL:X63382; NID:gl14170; PIDN:CAA44980.1; PID:gl14173
C/Genetics:
A/Gene: atpH

A/Genome: chloroplast
C/Superfamily: H+-transporting ATP synthase lipid-binding protein

C/Keywords: ATP biosynthesis; chloroplast; hydrolase; membrane-associated complex; thylakoid
F/7-33/Domain: transmembrane #status predicted <TM2>
F/51-77/Domain: transmembrane #status predicted <TM2>
F/61/Active site: Glu #status predicted

Query Match 2.8%; Score 6; DB 2; Length 82;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVALSL 144

DB 69 VVALSL 74

RESULT 94

S39516

H+-transporting two-sector ATPase (EC 3.6.3.14) lipid-binding protein - red alga (Cyanidium caldarium)
C/Species: chloroplast Cyanidium caldarium
C/Date: 07-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 03-Jun-2002

C/Accession: S39516

R/Kostrzewa, M.; Zetsche, K.

Plant Mol. Biol. 23, 67-76, 1993

A/Title: Organization of plastid-encoded ATPase genes and flanking regions including homeologs

A/Reference number: S39512; MUID:94033298; PMID:8219057

A/Accession: S39516

A/Status: translation not shown

A/Molecule type: DNA

A/Residues: 1-83 <KOS>

A/Cross-references: EMBL:X67814; NID:9429169; PIDN:CAA48021.1; PID:gl29174

A/Note: the source is designated as Galdieria sulphuraria

C/Genetics:

A/Gene: atpH

A/Genome: chloroplast

C/Superfamily: H+-transporting ATP synthase lipid-binding protein

C/Keywords: ATP biosynthesis; chloroplast; hydrolase; membrane-associated complex; thylakoid

Query Match 2.8%; Score 6; DB 2; Length 83;

Best Local Similarity 100.0%; Pred. No. 2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVALSL 144

DB 69 VVALSL 74

RESULT 95

S23424

H+-transporting two-sector ATPase (EC 3.6.3.14) lipid-binding protein - chromophytic alga
C/Species: chloroplast Pavlova lutheri
C/Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 03-Jun-2002

C/Accession: S23424

R/Scaramuzza, C.D.; Stokes, H.W.; Hiller, R.G.

FEBS Lett. 304, 119-123, 1992

A/Title: Characterisation of a chloroplast-encoded secY homologue and atpH from a chromophyte

A/Reference number: S23423; MUID:92316212; PMID:1618309

A/Accession: S23424

A/Molecule type: DNA

A/Residues: 1-83 <SCA>

A/Cross-references: EMBL:X64731; NID:gl2110; PIDN:CAA45997.1; PID:gl2112

C/Genetics:

A/Gene: atpH

A/Genome: chloroplast

C/Superfamily: H+-transporting ATP synthase lipid-binding protein

C/Keywords: ATP biosynthesis; chloroplast; hydrolase; membrane-associated complex; thylakoid

F/7-33/Domain: transmembrane #status predicted <TM1>

```

F:51-77/Domain: transmembrane #status predicted <TM2>
F:61/Active site: Glu #status predicted

Query Match      2.8%; Score 6; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVALSL 144
   |||||
Db 69 VVALSL 74

RESULT 96
D83389
conserved hypothetical protein PA2045 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: D83389
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bhandal, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: D83389
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-86 <STO>
A:Cross-references: GB:AE004631; GB:AE004091; NID:5948050; PIDN:AAG05433.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2045
C:Superfamily: conserved hypothetical protein HT1000

Query Match      2.8%; Score 6; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MKFLLI 7
   |||||
Db 1 MKFLLI 6

RESULT 97
H64540
hypothetical protein HP0168 - Helicobacter pylori (strain 26595)
C:Species: Helicobacter pylori
C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C:Accession: H64540
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A>Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: H64540
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-87 <FOM>
A:Cross-references: GB:AE000537; GB:AE000511; NID:92313247; PIDN:AD07241.1; PID:G231325

Query Match      2.8%; Score 6; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 LKQMKR 86
   |||||
Db 39 LKQMKR 44

RESULT 98
C71966
hypothetical protein jhp0154 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C:Accession: C71966
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999
A>Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: C71966
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-87 <ARN>
A:Cross-references: GB:AE001454; GB:AE001439; NID:94154666; PIDN:AAD05737.1; PID:G415467
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0154

Query Match      2.8%; Score 6; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 LKQMKR 86
   |||||
Db 39 LKQMKR 44

RESULT 99
C90667
hypothetical membrane protein [imported] - Escherichia coli (strain O157:H7, substrain RJ)
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: C90667
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: C90667
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-92 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA033730.1; PID:G13359764; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECS0307

Query Match      2.8%; Score 6; DB 2; Length 92;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVALSL 144
   |||||
Db 26 VVALSL 31

RESULT 100
G85517
hypothetical protein Z0344 [imported] - Escherichia coli (strain O157:H7, substrain EDL93)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: G85517
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, M.W.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; DiMalanta, E.; Potamocis, K.; Apodaca, Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: G85517
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-92 <STO>
A:Cross-references: GB:AE005174; NID:G12513061; PIDN:AAG54603.1; GSPDB:GN00145; UWGP:Z034

```


A;Experimental source: strain O157:H7, substrain EDL933
 C;Genetics:
 A;Gene: Z0344

Query Match 2.8%; Score 6; DB 2; Length 92;
 Best Local Similarity 100.0%; Pred.No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 139 VVALSL 144
 |||||
 Db 26 VVALSL 31

Search completed: August 6, 2004, 16:10:02
 Job time : 19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2004, 16:02:42 ; Search time 13 Seconds
(without alignments)
853.149 Million cell updates/sec

Title: US-10-024-955-7
Perfect score: 213
Sequence: 1 MMKELLIAAFAVAVSADPI.....VRKEMTKVLAPAKRSLKKN 213

Scoring table: OIIGO
Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	213	100.0	213	1	ALL7_DERPA
2	25	11.7	215	1	ALL7_DERPT
3	8	3.8	256	1	MORI_BUCAP
4	8	3.8	412	1	DHON_METGL
5	8	3.8	536	1	GAG_FSMVD
6	8	3.8	580	1	GAG_FLV
7	7	3.3	86	1	PTHF_STRMU
8	7	3.3	111	1	GDIR_CAVPO
9	7	3.3	149	1	YCL3_ARCFU
10	7	3.3	162	1	YX4D_RHISN
11	7	3.3	185	1	YB46_AQUAE
12	7	3.3	204	1	GDIR_BOVIN
13	7	3.3	204	1	GDIR_HUMAN
14	7	3.3	204	1	GDIR_MOUSE
15	7	3.3	208	1	YB23_CAEEL
16	7	3.3	252	1	YB25_VIBCH
17	7	3.3	255	1	Y166_METUA
18	7	3.3	295	1	CHIO_RAT
19	7	3.3	353	1	APJ_KENLA
20	7	3.3	354	1	OCB_AGRF5
21	7	3.3	356	1	OCB_AGRF4
22	7	3.3	430	1	SVH_CLOAB
23	7	3.3	434	1	Y166_CAEEL
24	7	3.3	437	1	PNCB_VIBVU
25	7	3.3	460	1	ORCS_DROME
26	7	3.3	468	1	CHIO_HUMAN
27	7	3.3	500	1	ARAA_YERPE
28	7	3.3	511	1	C4DL_DROME
29	7	3.3	556	1	CG48_SCHPO
30	7	3.3	833	1	VIRA_AGRF5
31	7	3.3	886	1	SVA_BABBA
32	7	3.3	2009	1	SEC7_YEAST
33	6	2.8	40	1	LHB1_ECTHL

RESULT 1

ALIGNMENTS

34	6	2.8	48	1	LHB1_RHOTE
35	6	2.8	52	1	LHB2_RHOTE
36	6	2.8	58	1	A95E_DROME
37	6	2.8	77	1	ACP_BACHD
38	6	2.8	82	1	ATPH_ANTSP
39	6	2.8	82	1	ATPH_CVACA
40	6	2.8	82	1	ATPH_FORPU
41	6	2.8	83	1	ATPH_GALSU
42	6	2.8	83	1	ATPH_PAVLU
43	6	2.8	86	1	YK45_PSEAE
44	6	2.8	88	1	PTPH_LACCA
45	6	2.8	94	1	ESXI_MYCBO
46	6	2.8	94	1	ESXI_MYCTU
47	6	2.8	101	1	NULC_PLEBO
48	6	2.8	117	1	Y13K_BPT4
49	6	2.8	124	1	GC5H_THEME
50	6	2.8	131	1	RS6_CHLTE
51	6	2.8	134	1	RUVK_HELPJ
52	6	2.8	134	1	RUVK_HELPY
53	6	2.8	135	1	JANA_DROME
54	6	2.8	146	1	HBEC_HOPLI
55	6	2.8	147	1	Y142_METJA
56	6	2.8	148	1	DUT_THETN
57	6	2.8	150	1	YMGW_YEAST
58	6	2.8	152	1	IL3_HYLLA
59	6	2.8	154	1	OM22_NEUCR
60	6	2.8	158	1	LUXS_LACLA
61	6	2.8	158	1	LUXS_LACPL
62	6	2.8	158	1	RNKD_SAGOE
63	6	2.8	158	1	SSRP_THETN
64	6	2.8	159	1	SSRP_ECOLI
65	6	2.8	159	1	SSRP_SALTY
66	6	2.8	160	1	ECF_GORGO
67	6	2.8	160	1	ECF_HUMAN
68	6	2.8	160	1	ECF_PANTR
69	6	2.8	160	1	ECF_PONPY
70	6	2.8	160	1	SSRP_HAEDU
71	6	2.8	160	1	SSRP_PHOLL
72	6	2.8	160	1	YQGC_BACSU
73	6	2.8	161	1	RNKD_GORGO
74	6	2.8	161	1	RNKD_HUMAN
75	6	2.8	161	1	RNKD_PANTR
76	6	2.8	161	1	RNKD_PONPY
77	6	2.8	162	1	NOS2_MACMU
78	6	2.8	170	1	FMCI_ECOLI
79	6	2.8	172	1	APT_ANASP
80	6	2.8	174	1	YA68_METTH
81	6	2.8	176	1	PHB8_AGLNE
82	6	2.8	178	1	Y045_CAUCR
83	6	2.8	181	1	APT2_YEAST
84	6	2.8	185	1	DSBE_SALTY
85	6	2.8	185	1	DSBE_SALTY
86	6	2.8	186	1	DYR_HUMAN
87	6	2.8	192	1	DENR_CAEEL
88	6	2.8	197	1	TSAA_BUCAL
89	6	2.8	199	1	Y5E0_VIBPA
90	6	2.8	200	1	LEUD_ECOS7
91	6	2.8	200	1	LEUD_ECOLI
92	6	2.8	207	1	ENO_CAMPE
93	6	2.8	208	1	EFH2_TRYCR
94	6	2.8	209	1	UPP_OCEIH
95	6	2.8	209	1	UPP_RHIME
96	6	2.8	215	1	YPCV_BACSU
97	6	2.8	216	1	PROZ_XANCG
98	6	2.8	217	1	HRCR_RALSO
99	6	2.8	219	1	YQIY_BACSU
100	6	2.8	226	1	NHB2_RRORH

P80590	rhodocyclu
P80591	rhodocyclu
P16548	drosophila
Q9ka04	bacillus ha
Q02851	antithamio
Q9tm30	cyanidium c
P51246	porphyra pu
P35013	gaideria s
P28530	pavlova lut
Q91270	pseudomonas
Q9k1v3	lactobacill
P59802	mycobacteri
P96314	mycobacteri
Q00244	plectonema
P39504	bacterioph
Q9wy55	thermotoga
Q8kam1	chlorobium
Q9zma5	helicobacte
Q25101	helicobacte
P20348	drosophila
P82316	hoplosternu
Q57607	methanococc
Q8ra46	thermoanaer
Q03579	saccharomyc
P06740	hylobates 1
Q07335	neurospora
Q9ciu0	lactococcus
Q89y16	lactobacill
P47786	saginus oe
Q8zb39	thermoanaer
P32052	escherichia
P43658	salmonella
P47778	gorilla gor
P12724	homo sapien
P47780	pan troglod
P47781	pongo pygma
Q7vm64	haemophilus
Q7niul	phototroba
P54486	bacillus su
P47782	gorilla gor
P10153	homo sapien
P47785	pan troglod
P47784	pongo pygma
O46660	macaca mula
P02971	escherichia
Q8yni3	anabaena sp
O27140	methanobact
P28560	aglaothami
Q9ac22	caulobacter
P36973	saccharomyc
Q8xfk6	salmonella
Q8xfk6	salmonella
P00374	homo sapien
Q9nah4	caenorhabdi
P57279	buchnera ap
Q87jv8	vibrio para
Q8xa01	escherichia
P30126	escherichia
O30885	campylobact
P41049	trypanosoma
Q8em74	oceanobacil
Q92c49	rhizobium m
P42978	bacillus su
P37828	xanthomonas
Q52488	raistonia s
P54536	bacillus su
P29379	rhodococcus

ALL7_DERPT
ID ALL7_DERPT STANDARD; PRT; 213 AA.
AC Q26456;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Mite allergen Der f 7 precursor (Der f VII).
GN DERF7.
OS Dermatophagoides farinae (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcophormes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6954;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96120794; PubMed=8856554;
RX Shen H.-D., Chua K.-Y., Lin W.L., Hsieh K.-H., Thomas W.R.;
RT "Molecular cloning and immunological characterization of the house
dust mite allergen Der f 7";
RL Clin. Exp. Allergy 25:1000-1006(1995).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the mite group 7 allergen family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; S80655; AAB35977.1; -.
KW Allergen; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 213 MITE ALLERGEN DER F 7.
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 213 AA; 23627 MW; 3CF1F529107B7808 CRC64;
Query Match 100.0%; Score 213; DB 1; Length 213;
Best Local Similarity 100.0%; Pred. No. 2.7e-211; Mismatches 0; Indels 0; Gaps 0;
Matches 213; Conservative 0;
QY 1 MKKFLIIAAVAVFVAVSADPIHYDKITEINKAIDDAIAIEQSEITIDPMKVPDCHADKFER 60
Db 1 MKKFLIIAAVAVFVAVSADPIHYDKITEINKAIDDAIAIEQSEITIDPMKVPDCHADKFER 60
QY 61 HVGIVDFKGLAMRNIEARGLKQKQGDANVKGEIGVKAHLIGVHDDIVSMYDILAY 120
Db 61 HVGIVDFKGLAMRNIEARGLKQKQGDANVKGEIGVKAHLIGVHDDIVSMYDILAY 120
QY 121 KLGLDHPHTHVSIDQDFVVALSLSEISDEGNITMTSFVRQFANVNVHIGGLSILDPFG 180
Db 121 KLGLDHPHTHVSIDQDFVVALSLSEISDEGNITMTSFVRQFANVNVHIGGLSILDPFG 180
QY 181 VLSVDLTAIFQDTVRKEMTKVLAFAKFELEKN 213
Db 181 VLSVDLTAIFQDTVRKEMTKVLAFAKFELEKN 213
RESULT 2
ALL7_DERPT
ID ALL7_DERPT STANDARD; PRT; 215 AA.
AC P49273;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Mite allergen Der p 7 precursor (Der p VII).
GN DERP7.
OS Dermatophagoides pteronyssinus (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcophormes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.

OX NCBI_TaxID=6956;
RN [1]
RP SEQUENCE FROM N.A.
RA Shen H.-D., Chua K.-Y., Lin K.-L., Hsieh K.-H., Thomas W.R.;
RT "Molecular cloning of a house dust mite allergen with common antibody
binding specificities with multiple components in mite extracts";
RL Clin. Exp. Allergy 23:934-940(1993).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the mite group 7 allergen family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U37044; AAB80264.1; -.
KW Allergen; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 215 MITE ALLERGEN DER P 7.
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 215 AA; 23881 MW; 63AD03DB5C8CB1C0 CRC64;
Query Match 11.7%; Score 25; DB 1; Length 215;
Best Local Similarity 100.0%; Pred. No. 5.7e-18; Mismatches 0; Indels 0; Gaps 0;
Matches 25; Conservative 0;
QY 155 TSFEVRQFANVNVHIGGLSILDPFG 179
Db 155 TSFEVRQFANVNVHIGGLSILDPFG 179
RESULT 3
MURI_BUCAP
ID MURI_BUCAP STANDARD; PRT; 256 AA.
AC Q8K924;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Glutamate racemase (EC 5.1.1.3).
GN MURI OR BUSQ536.
OS Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=98794;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=22084549; PubMed=12089438;
RA Tamás I., Klasson L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,
RA Wernegreen J.J., Sandstroem J.P., Moran N.A., Andersson S.G.E.;
RT "50 million years of genomic stasis in endosymbiotic bacteria";
RL Science 296:2376-2379(2002).
CC -!- FUNCTION: Provides the (R)-glutamate required for cell wall
biosynthesis (by similarity).
CC -!- CATALYTIC ACTIVITY: L-glutamate = D-glutamate.
CC -!- PATHWAY: Peptidoglycan biosynthesis.
CC -!- SIMILARITY: Belongs to the aspartate/glutamate racemases family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AE014127; AAM68077.1; -.
DR HMAP; MF_00258; -; 1.
DR InterPro; IPR004391; Glu_race.

```
DR Pfam; PF01177; Asp Glu race; 1.
DR TIGRPFAMs; TIGR00067; Glut race; 1.
DR PROSITE; PS00923; ASP_GLU_RACEMASE_1; 1.
DR PROSITE; PS00924; ASP_GLU_RACEMASE_2; 1.
KW Peptidoglycan synthesis; Cell wall; isomerase; Complete proteome.
SQ SEQUENCE 256 AA; 29628 MW; F8428EBECF3059B2 CRC64;

Query Match 3.8%; Score 8; DB 1; Length 256;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 169 IGLSILD 176
Db 8 IGLSILD 15

RESULT 4
DHON METGL STANDARD; PRT; 412 AA.
AC P37144;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Homoserine dehydrogenase [EC 1.1.1.3] (Hdh).
GN HOM.
OS Methylobacillus glycogenes.
OC Bacteria; Proteobacteria; Betaproteobacteria; Methylobacterales;
OC Methylobacteriaceae; Methylobacillus.
OX NCBI_TaxID=406;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 213171;
RX MEDLINE=94161493; PubMed=8117070;
RA Motoyama H., Maki K., Anazawa H., Ishino S., Teshiba S.;
RT "Cloning and nucleotide sequences of the homoserine dehydrogenase
RT genes (hom) and the threonine synthase genes (thrc) of the Gram-
RT negative obligate methylophilic Methylobacillus glycogenes.";
RL Appl. Environ. Microbiol. 60:111-119 (1994).
CC -1- CATALYTIC ACTIVITY: L-homoserine + NAD(P) (+) = L-aspartate 4-
CC semialdehyde + NAD(P)H.
CC -1- PATHWAY: Conversion of L-aspartate to homoserine; third step.
CC -1- SIMILARITY: Belongs to the homoserine dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D14071; BAA40415.1; -.
DR InterPro; IPR002912; ACT.
DR InterPro; IPR001342; Homoserine dh.
DR InterPro; IPR005106; NAD_binding_3.
DR Pfam; PF01842; ACT; 1.
DR Pfam; PF00742; Homoserine dh; 1.
DR Pfam; PF03447; NAD binding 3; 1.
DR PROSITE; PS01042; HOMOSER_DHGNASE; 1.
KW Oxidoreductase; NADP; Threonine biosynthesis; Isoleucine biosynthesis;
KW Methionine biosynthesis.
FT NP BIND 9 16 NADP (POTENTIAL).
SQ SEQUENCE 412 AA; 44818 MW; 613A1B7FDECF4AA CRC64;

Query Match 3.8%; Score 8; DB 1; Length 412;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 DDAAIAIE 41
Db 386 DDAAIAIE 393

us-10-024-955-7.oligo.rsp
```

```
RESULT 5
GAG FSVMD
ID - GAG_FSVMD STANDARD; PRT; 536 AA.
AC P03340;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE GAG polyprotein [Contains: Core protein P15; Core protein P12; Core
DE protein P30; Core protein P10].
GN GAG.
OS Feline sarcoma virus (strain McDonough).
OC Viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses.
OX NCBI_TaxID=11778;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84119469; PubMed=6582485;
RA Hampe A., Gobet M., Sherr C.J., Galibert F.;
RT "Nucleotide sequence of the feline retroviral oncogene v-fms shows
RT unexpected homology with oncogenes encoding tyrosine-specific protein
RT kinases.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:95-99 (1984).
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-FMS
CC POLYPROTEIN.
CC -1- SIMILARITY: TO MOUSE TESTOSTERONE-REGULATED RP2 PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K01643; AAA3045.1; ALT_TERM.
DR FIRM; A03938; FOMVMD.
DR InterPro; IPR000840; Gag MA.
DR InterPro; IPR002079; Gag p12.
DR InterPro; IPR003036; Gag p30.
DR Pfam; PF01140; Gag MA; 1.
DR Pfam; PF01141; Gag p12; 1.
DR Pfam; PF02093; Gag p30; 1.
KW Core protein; Polyprotein; Myristate; Lipoprotein.
FT PROPEP 1 77 LEADER PEPTIDE.
FT CHAIN 78 204 CORE PROTEIN P15.
FT CHAIN 205 274 CORE PROTEIN P12.
FT CHAIN 275 522 CORE PROTEIN P30.
FT CHAIN 523 536 CORE PROTEIN P10.
FT LIPID 79 79 N-myristoyl glycine (in host).
SQ SEQUENCE 536 AA; 60234 MW; FBBC80D612AC8702 CRC64;

Query Match 3.8%; Score 8; DB 1; Length 536;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 196 KEMTKVLA 203
Db 516 KEMTKVLA 523

RESULT 6
GAG_FLV
ID - GAG_FLV STANDARD; PRT; 580 AA.
AC P10262; Q85560;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE GAG polyprotein [Contains: Core protein P15; Core protein P12; Core
DE protein P30; Core protein P10].
GN GAG.
OS Feline leukemia virus.
OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus.
OX NCBI_TaxID=11768;
```

```
RN SEQUENCE FROM N.A.
RX MEDLINE=84216470; PubMed=6328019;
RA Laprevotte I., Hampe A., Sherr C.J., Galibert F.;
RT "Nucleotide sequence of the gag gene and gag-pol junction of feline
RL leukemia virus";
RL J. Virol. 50:884-894 (1984).
CC -!- SIMILARITY: Contains 1 CCHC-type zinc finger.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K01803; AAA43054.1; -
DR EMBL; K01803; AAA43055.1; ALT INIT.
DR InterPro: IPR000840; Gag_MA.
DR InterPro: IPR002079; Gag_P12.
DR InterPro: IPR003036; Gag_P30.
DR InterPro: IPR001878; Znf_CCHC.
DR Pfam; PF01140; Gag_MA; 1.
DR Pfam; PF02093; Gag_P12; 1.
DR Pfam; PF02093; Gag_P30; 1.
DR Pfam; PF00098; zf_CCHC; 1.
DR PRINTS; PR00939; C2HCZNRINGER.
DR SMART; SM00343; ZNF_C2HC; 1.
DR PROSITE; PS0158; ZF_CCHC; 1.
KW Core protein; Polyprotein; Zinc-finger.
FT PROPEP 1 75 LEADER PEPTIDE.
FT CHAIN 76 201 CORE PROTEIN P15.
FT CHAIN 202 271 CORE PROTEIN P12.
FT CHAIN 272 519 CORE PROTEIN P30.
FT CHAIN 520 580 CORE PROTEIN P10.
FT ZNFING 547 584 CCHC-TYPE.
SQ SEQUENCE 580 AA; 65195 MW; 10F8C375B37042C CRC64;

Query Match 3.8%; Score 8; DB 1; Length 580;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 KEMTKVLA 203
DB 513 KEMTKVLA 520

RESULT 7
ID ETHP STRMU STANDARD; PRT; 86 AA.
AC P45596;
DT 01-NOV-1995 (Rel. 32, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphocarrier protein HPr (Histidine-containing protein).
GN PTSH OR SMU.674.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NG5 / Serotype C;
RA MEDLINE=94178918; PubMed=8132321;
RA Boyd D.A., Cvitkovitch D.G., Hamilton I.R.;
RT "Sequence and expression of the genes for Hpr (ptsh) and enzyme I
RT (ptei) of the phosphoenolpyruvate-dependent phosphotransferase
RL transport system from Streptococcus mutans.";
RL Infect. Immun. 62:1156-1165 (1994).
RN [2]
RP SEQUENCE.
RC STRAIN=Ingbritt;
```

```
RX MEDLINE=94197719; PubMed=8147873;
RA Dashper S.G., Kirsbaum L., Huq N.L., Riley P.F., Reynolds E.C.;
RT "Complete amino acid sequence and comparative molecular modelling of
RT HPr from Streptococcus mutans Ingbritt.";
RL Biochem. Biophys. Res. Commun. 199:1297-1304 (1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=2225063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar P., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen.";
RC Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
CC -!- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOENOLPYRUVATE-DEPENDENT
CC SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYDRATE ACTIVE
CC -TRANSPORT SYSTEM. THE PHOSPHORYL GROUP FROM PHOSPHOENOLPYRUVATE
CC (PEP) IS TRANSFERRED TO THE PHOSPHORYL CARRIER PROTEIN HPR BY
CC ENZYME I. PHOSPHO-HPR THEN TRANSFERS IT TO THE PERMEASE (ENZYMES
CC II/III). HPR IS COMMON TO ALL PTS.
CC -!- ENZYME REGULATION: PHOSPHORYLATION ON SER-45 INHIBITS THE
CC PHOSPHORYL TRANSFER FROM ENZYME I TO HPR (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the Hpr family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; LJ5191; AAA91092.1; -
DR EMBL; AJ014911; AAN58408.1; -
DR PIR; A44562; A44562.
DR HSP; P07515; 1PTF.
DR InterPro: IPR001020; HPr_Hisp_S.
DR InterPro: IPR000032; HPr_protein.
DR InterPro: IPR002114; HPr_Serp_S.
DR InterPro: IPR005698; PTS_HPr.
DR Pfam; PF00381; PTS-HPr; 1.
DR PRINTS; PR00107; PHOSPHOCHPR.
DR ProDom; PD00238; HPr_protein; 1.
DR TIGRFAMs; TIGR01003; PTS_HPr_family; 1.
DR PROSITE; PS00369; PTS_HPR_HIS; 1.
DR PROSITE; PS00589; PTS_HPR_SER; 1.
KW Phosphotransferase system; Sugar transport; Phosphorylation;
KW Complete proteome.
FT INIT_MET 0 0 PARTIAL.
FT MOD_RES 14 14 PHOSPHORYLATION (BY ENZYME I) (BY
FT SIMILARITY).
FT MOD_RES 45 45 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 86 AA; 8805 MW; A57B9CD819341A64 CRC64;

Query Match 3.3%; Score 7; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 DDAAIAAI 40
DB 70 DDAAIAAI 76

RESULT 8
ID GDPR_CAVPO STANDARD; PRT; 111 AA.
AC P80237;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Rho GDP-dissociation inhibitor 1 (Rho GDI 1) (Rho-GDI alpha)
```

```

DE (Fragments).
GN ARHGDA.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE.
RC STRAIN=Hartley; TISSUE=Macrophage;
RX MEDLINE=94039069; PubMed=8223583;
RA Pick E.; Gorzalczyk Y.; Engel S.;
RT "Role of the rac1 p21-GDP-dissociation inhibitor for rho heterodimer
RT in the activation of the superoxide-forming NADPH oxidase of
RT macrophages."
RL Eur. J. Biochem. 217:441-455(1993).
CC -1- FUNCTION: Regulates the GDP/GTP exchange reaction of the Rho
CC proteins by inhibiting the dissociation of GDP from them, and the
CC subsequent binding of GTP to them.
CC -1- FUNCTION: Rac1 p21/rho GDI heterodimer is the active component of
CC the cytosolic factor sigma 1, which is involved in stimulation of
CC the NADPH oxidase activity in macrophages.
CC -1- SUBUNIT: FORMS A HETERODIMER WITH P21/RAC-1.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- PM: The N-terminus is blocked.
CC -1- SIMILARITY: Belongs to the Rho GDI family.
DR HSP; P19803; IGDF.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR000406; Rho_GDI.
DR PRINTS; PR00492; RHOGDI.
KW GTPase activation.
FT NON_TER 1 1
FT NON_CONS 40 41
FT NON_CONS 46 47
FT NON_CONS 86 87
FT UNSURE 98 98
FT UNSURE 100 100
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 12496 MW; 0FCE35BEA40FE951 CRC64;

Query Match 3.3%; Score 7; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSADP 19
DB 1 VAVSADP 7
|||||

RESULT 9
ID_YC13_ARCFU STANDARD; PRT; 149 AA.
AC Q25055;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF1213.
GN AF1213.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.B., Kyripides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,

```

```

RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001020; AAB90034.1; -
DR PIR; D69401; D69401.
DR TIGR; AF1213; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 149 AA; 17778 MW; 955AE6E28093F77F CRC64;

Query Match 3.3%; Score 7; DB 1; Length 149;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 FKRELEK 212
DB 7 FKRELEK 13
|||||

RESULT 10
ID_Y4XD_RHISN STANDARD; PRT; 162 AA.
AC P55695;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 18.0 kDa protein Y4XD.
GN Y4XD.
OS Rhizobium sp. (strain NGR234).
OG Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97305956; PubMed=9163424;
RX Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -1- SIMILARITY: TO SIMILAR PROTEINS IN OTHER NITROGEN-FIXING BACTERIA.
CC THIS PROTEIN IS GENERALLY FOUND IN THE NIFX-NIFW INTERGENIC
CC REGION.
CC -1- SIMILARITY: STRONG, TO Y4VQ.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000105; AAB91926.1; -
DR PIR; T10831; T10831.
DR InterPro; IPR004952; DUF269.
DR Pfam; PF03270; DUF269; 1.
DR ProDom; PD008304; DUF269; 1.
KW Hypothetical protein; Nitrogen fixation; Plasmid.
SQ SEQUENCE 162 AA; 18015 MW; 0B4D5C0FCB2BB61 CRC64;

Query Match 3.3%; Score 7; DB 1; Length 162;

```

Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 LEISDEG 150
|||||
Db 100 LEISDEG 106

RESULT 11

YE46 AQUAE
ID YE46 AQUAE STANDARD; PRT; 185 AA.
AC G67433;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AQ_1446.
GN AQ_1446.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9337320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujaay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus.";
RL Nature 392:353-358 (1998).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: STRONG, TO A.AEOLICUS AQ 1900.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE000741; AAC07402.1; --
DR PIR: G70425; G70425.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 1 21
FT TRANSMEM 111 131 POTENTIAL.
SQ SEQUENCE 185 AA; 22670 MW; EDA145E48ED739C9 CRC64;

Query Match 3.3%; Score 7; DB 1; Length 185;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMKFLLI 7
|||||
Db 1 MMKFLLI 7

RESULT 12

GDIR BOVIN
ID GDIR BOVIN STANDARD; PRT; 204 AA.
AC P19803;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Rho GDP-dissociation inhibitor 1 (Rho GDI 1) (Rho-GDI alpha).
GN ARGGDI.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC TISSUE=Brain;
RX MEDLINE=91016432; PubMed=2120668;
RA Fukumoto Y., Kaibuchi K., Hori Y., Fujioka H., Araki S., Ueda T.,
RA Kikuchi A., Takai Y.;
RT "Molecular cloning and characterization of a novel type of regulatory
RT protein (GDI) for the rho proteins, ras p21-like small GTP-binding
RT proteins.";
RL Oncogene 5:1321-1328 (1990).
RN [2]
RP STRUCTURE BY NMR OF 60-204.
RX MEDLINE=97337869; PubMed=9194563;
RA Gesser Y.O., Nomanbhoy T.K., Aghazadeh B., Manor D., Combs C.,
RA Cerione R.A., Rosen M.K.;
RT "C-terminal binding domain of Rho GDP-dissociation inhibitor directs
RT N-terminal inhibitory peptide to GTPases.";
RL Nature 387:814-819 (1997).
CC -!- FUNCTION: Regulates the GDP/GTP exchange reaction of the Rho
CC proteins by inhibiting the dissociation of GDP from them, and the
CC subsequent binding of GTP to them.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Brain, lung, thymus, spleen, small intestine,
CC and kidney, and weakly in heart and liver.
CC -!- SIMILARITY: Belongs to the Rho GDI family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements/>
CC or send an email to license@isb-sib.ch).
CC

EMBL: X52689; CAA36916.1; --
DR PIR: S12121; S12121.
DR PDB: 1G3F; 19-NOV-97.
DR PDB: 1A0W; 19-NOV-97.
DR PDB: 1DOA; 09-FEB-00.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR000406; Rho GDI.
DR Pfam: PF02115; Rho GDI; I.
DR PRINTS: PR00492; RHOGDI.
KW GTPase activation; Phosphorylation; 3D-structure.
FT DOMAIN 66 83 HYDROPHOBIC.
FT MOD RES 101 101 PHOSPHORYLATION (BY PKA) (POTENTIAL).
FT MOD RES 115 115 PHOSPHORYLATION (BY PKC) (POTENTIAL).
FT STRAND 70 78
FT STRAND 88 89
FT STRAND 95 95
FT TURN 96 99
FT HELIX 101 103
FT STRAND 106 107
FT TURN 109 118
FT STRAND 123 133
FT STRAND 138 149
FT STRAND 155 157
FT STRAND 160 164
FT TURN 169 171
FT STRAND 176 182
FT STRAND 188 197
FT TURN 202 204
SQ SEQUENCE 204 AA; 23421 MW; 49CE7DEB05D271CA CRC64;

Query Match 3.3%; Score 7; DB 1; Length 204;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSADP 19
|||||
Db 59 VAVSADP 65

RESULT 13

GDH_HUMAN
ID_GDIH_HUMAN STANDARD; PRT; 204 AA.
AC P52565;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Rho GTP-dissociation inhibitor 1 (Rho GDI 1) (Rho-GDI alpha).
GN ARHGAP1 OR GDI1
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94085490; PubMed=8262133;
RA Leffers H., Nielsen M.S., Andersen A.H., Honore B., Madsen P.,
RA Vandekeckevoort J., Celis J.E.;
RT "Identification of two human Rho GDP dissociation inhibitor proteins
RT whose overexpression leads to disruption of the actin cytoskeleton";
RL Exp. Cell Res. 209:165-174(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Maeda A., Kaibuchi K., Takai Y.;
RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Chuang T.H., Bokoch G.M.;
RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Mulholland J.G., Schwinn D.A., Caron M.G., Liggett S.B.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA TISSUE=Brain;
RC "cDNA clones of human proteins involved in signal transduction
RT sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA TISSUE=Colon, Lung, Muscle, Skin, Tonsil, and Uterus;
RC MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buote K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T.I., Wang J., Haie H.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield V.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 59-204.
RX MEDLINE=97341226; PubMed=9195862;
RA Keep N.C.E., Barnes M., Barsukov I., Badii R., Lian L.-Y., Segal A.W.,
RA Moody P.C.E., Roberts G.C.K.;
RT "A modulator of rho family G proteins, rhoGDI, binds these G proteins
RT via an immunoglobulin-like domain and a flexible N-terminal arm";
RL Structure 5:623-633(1997).
CC -!- FUNCTION: Regulates the GTP/GDP exchange reaction of the Rho
CC proteins by inhibiting the dissociation of GDP from them, and the

CC subsequent binding of GTP to them (By similarity).
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the Rho GDI family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X69550; CAA49281.1; -
CC EMBL; D13989; BAA03096.1; -
CC EMBL; M7579; AAR36566.1; -
CC EMBL; X63863; CAA45344.1; -
CC EMBL; AF498926; AAM21074.1; -
CC EMBL; BC005851; AAH05851.1; -
CC EMBL; BC005875; AAH05875.1; -
CC EMBL; BC008701; AAH08701.1; -
CC EMBL; BC009759; AAH09759.1; -
CC EMBL; BC016031; AAH16031.1; -
CC EMBL; BC016185; AAH16185.1; -
CC EMBL; BC024258; AAH24258.1; -
CC EMBL; BC027730; AAH27730.1; -
CC F1R; I38156; I38156.
CC PDB; 1RHO; 15-OCT-97.
CC PDB; 1F50; 02-MAY-01.
CC PDB; 1F5T; 02-MAY-01.
CC PDB; 1F70; 02-MAY-01.
CC PDB; 1F73; 02-MAY-01.
CC PDB; 1HH4; 28-AUG-01.
CC PDB; 1KMT; 11-DEC-02.
CC Aarhus/Ghent-2DPAGE; 8118; IEF.
CC Genew; HGNC:678; ARHGDI.A.
CC MIM; 601925; -
CC GO; GO:0005515; F:protein binding; TAS.
CC GO; GO:0005094; F:rho GDP-dissociation inhibitor activity; TAS.
CC GO; GO:0007162; P:negative regulation of cell adhesion; TAS.
CC GO; GO:0007266; P:rho protein signal transduction; TAS.
CC InterPro; IPR007110; IG-like.
CC InterPro; IPR000406; Rho GDI.
CC Pfam; PF02115; Rho GDI; I.
CC PRINTS; PR00482; RHOGDI.
KW GTPase activation; 3D-structure. I -> V (IN REF. 3).
FT CONFLICT 139 139
FT STRAND 70 78
FT TURN 80 81
FT STRAND 87 89
FT TURN 90 91
FT HELIX 95 99
FT STRAND 102 105
FT TURN 106 107
FT STRAND 109 118
FT STRAND 123 134
FT TURN 135 136
FT STRAND 137 149
FT STRAND 156 159
FT STRAND 163 184
FT TURN 169 171
FT STRAND 173 182
FT TURN 184 185
FT STRAND 190 199
SQ SEQUENCE 204 AA; 23207 MW; 59CB6F42E3B3BCCA CRC64;
Query Match 3.3%; Score 7; DB 1; Length 204;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 VAVSADP 19
DB 59 VAVSADP 65

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
RESULT 14
GDRI MOUSE
ID Q99PT1; Q99KC4; STANDARD; PRT; 204 AA.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Rho GDP-dissociation inhibitor 1 (Rho GDI 1) (Rho-GDI alpha) (GDI-1).
GN ARHGDI1 OR GDI1 OR CB7222.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP Minamitani T., Matsumoto K.;
RA "Mouse cDNA sequence for RhoGDI-1."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.L., Fellings E.A., Grouse L.H., Derge J.G.,
RA Strausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusana K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Regulates the GDP/GTP exchange reaction of the Rho
CC proteins by inhibiting the dissociation of GDP from them, and the
CC subsequent binding of GTP to them (By similarity).
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the Rho GDI family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; AB055070; BAB21527.1; -
CC EMBL; BC004732; AAH04732.1; -
CC HSSP; FI9803; 1GDF.
CC PMMA-2DPAGE; Q99PT1; -
CC MGD; MGI:2178103; Arhgdia.
CC InterPro; IPR007110; IG-like.
CC Pfam; PF02115; Rho GDI; I.
CC PRINTS; PR00492; RHOGDI.
CC GTPase activation.
CC CONFLICT 55 55 L -> P (IN REF. 2).
CC SEQUENCE 204 AA; 23407 MW; 8ACE6F4456D842D8 CRC64;
Query Match 3.3%; Score 7; DB 1; Length 204;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
RESULT 15
YS23_CAEEL
ID YS23_CAEEL STANDARD; PRT; 208 AA.
AC Q99365;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 23.3 kDa protein ZK1320.3 in chromosome II.
GN ZK1320.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RC Berks M.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: STRONG, TO C.ELEGANS ZK1320.2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; Z46934; CAA87041.1; -
CC PIR; T27749; T27749.
CC WormPep; ZK1320.3; CE01701.
KW Hypothetical protein.
SQ SEQUENCE 208 AA; 23306 MW; 356A84E57A6C163B CRC64;
Query Match 3.3%; Score 7; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
RESULT 16
YC85_VIBCH
ID YC85_VIBCH STANDARD; PRT; 252 AA.
AC Q9KSH1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0249 protein VCL1285.
GN VCL1285.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
SEQUENCE FROM N.A.
RP STEIN-EL Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
```

RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
RL *cholerae*;"
RW Nature 406:477-483(2000).
CC -!- SIMILARITY: Belongs to the UPF0249 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE004207; AAP94444.1; -
DR PIR; H82219; H82219.
DR TIGR; VC1285; -
DR HAMAP; MF 01246; -; 1.
DR InterPro; IPR006879; YdJc.
DR Pfam; PF04794; YdJc; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 252 AA; 28663 MW; 5B5E2C734D826B30 CRC64;

Query Match 3.3%; Score 7; DB 1; Length 252;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 GIVKAHL 103
DB 19 GIVKAHL 25
|||||

RESULT 17
Y166 METJA STANDARD; PRT; 255 AA.
AC Q57630;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein MJ0166.
GN MJ0166.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness B.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenn H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, *Methanococcus*
RL *jannaschii*." Science 273:1058-1073(1996).
CC -!- SIMILARITY: Belongs to the UPF0204 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U67473; BAB98148.1; -
DR PIR; G64320; G64320.
DR TIGR; MJ0166; -

DR HAMAP; MF 00562; -; 1.
DR InterPro; IPR007508; DUF516.
DR Pfam; PF04414; DUF516; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 255 AA; 29072 MW; 6428DF7AEC802CE4 CRC64;

Query Match 3.3%; Score 7; DB 1; Length 255;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MKFLLIA 8
DB 1 MKFLLIA 7
|||||

RESULT 18
CHIO RAT STANDARD; PRT; 295 AA.
AC Q03070;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Beta-chimaerin (Beta-chimerin).
GN CHN2 OR BCH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=Wistar; TISSUE=Testis;
RX MEDLINE=93179371; PubMed=8440677;
RA Leung T., How B.E., Manser E., Lim L.;
RT "Germ cell beta-chimaerin, a new GTPase-activating protein for
RT p21rac, is specifically expressed during the acrosomal assembly stage
RT in rat testis." J. Biol. Chem. 268:3813-3816(1993).
RL J. Biol. Chem. 268:3813-3816(1993).
CC -!- FUNCTION: GTPase activating protein for p21-rac.
CC -!- SUBCELLULAR LOCATION: Membrane-associated (potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Beta-1;
CC IsoId=Q03070-1; Sequences=Displayed;
CC Name=Beta-2;
CC IsoId=Q03070-2; Sequences=Not described;
CC -!- TISSUE SPECIFICITY: Found in cerebellum and testis.
CC -!- DEVELOPMENTAL STAGE: Expressed specifically in late stage
CC spermatocytes. In the cerebellum, emergence of beta-2 isoform
CC coincides with granule cells maturation and exhibits postnatal
CC developmental increases. Expression is specifically reduced in
CC weaver mutant.
CC -!- SIMILARITY: Contains 1 Rho-GAP domain.
CC -!- SIMILARITY: Contains 1 zinc-dependent phorbol-ester and DAG
CC binding domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X69489; CAA49244.1; -
DR EMBL; L07434; AAA40809.1; -
DR PIR; A45485; S29956.
DR HSSP; P28867; IPTQ.
DR InterPro; IPR002219; DAG PE-bind.
DR InterPro; IPR008936; Rho GAP.
DR InterPro; IPR000198; RhoGAP.
DR Pfam; PF00130; DAG PE-bind; 1.
DR Pfam; PF00820; RhoGAP; 1.
DR PRINTS; PR00008; DAGPEOVAIN.

DR SMART; SM00109; CL; 1.
 DR SMART; SM00324; Rhogap; 1.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
 DR PROSITE; PS0081; DAG_PE_BIND_DOM_2; 1.
 DR PROSITE; PS50238; Rhogap; 1.
 KW GTPase activation; Phorbol-ester binding; zinc; Membrane; SH2 domain;
 KW Alternative splicing.
 FT DOMAIN 42 91 PHORBOL-ESTER AND DAG BINDING.
 FT DOMAIN 104 295 RHO-GAP
 FT SEQUENCE 295 AA; 33837 MW; D7692D957B4816BD CRC64;
 SQ SEQUENCE 295 AA; 33837 MW; D7692D957B4816BD CRC64;
 Query Match 3.3%; Score 7; DB 1; Length 295;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 76 IEARGLK 82
 DB 128 IEARGLK 134
 RESULT 19
 APJ_XENLA STANDARD; PRT; 353 AA.
 AC P79560; P70058; (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE G protein-coupled receptor APJ homolog (Angiotensin receptor related
 DE protein) (Mesenchyme-associated serpentine receptor).
 GN X-MSR OR XANGIO1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97109511; PubMed=8951791;
 RA Devic E., Paqueriau L., Vernier P., Knibiehler B., Audigier Y.;
 RT "Expression of a new G protein-coupled receptor X-msr is associated
 RT with an endothelial lineage in Xenopus laevis."
 RL Mech. Dev. 59:129-140(1996).
 RN [2]
 RP SEQUENCE OF 1-303 FROM N.A.
 RA Saha M.S., Oakes J.A., Miles R.R.;
 RT "Xangio1, a novel Xenopus gene, is expressed in vascular precursor
 RT cells."
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Putative receptor for an apelin-like peptide coupled to
 CC G proteins that inhibit adenylate cyclase activity.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: At the gastrula stage, exclusively expressed
 CC in the mesodermal layer and at the neurula stage in the lateral
 CC plate mesoderm. Larval expression is observed in the endothelium
 CC of the primary blood vessels and the forming heart.
 CC -!- DEVELOPMENTAL STAGE: First expressed at the late blastula stage,
 CC increases during gastrulation and remains constant between neurula
 CC and larva stages.
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X93045; CAA63612.1; -.
 CC EMBL; U72029; AAB17004.1; -.
 CC InterPro; IPR000276; GPCR_Rhodpsn.
 CC Pfam; PF00001; 7tm 1; 1.
 CC PRINTS; PR00237; GFCRRHODPSN.

DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
 KW G-protein coupled receptor; Glycoprotein; Transmembrane.
 FT DOMAIN 1 38 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 39 61 1 (POTENTIAL).
 FT DOMAIN 62 72 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 73 95 2 (POTENTIAL).
 FT DOMAIN 96 109 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 110 131 3 (POTENTIAL).
 FT DOMAIN 132 150 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 151 173 4 (POTENTIAL).
 FT DOMAIN 174 206 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 207 229 5 (POTENTIAL).
 FT DOMAIN 230 248 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 249 270 6 (POTENTIAL).
 FT DOMAIN 271 294 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 295 317 7 (POTENTIAL).
 FT DOMAIN 318 353 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 19 19 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 181 181 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 150 150 P -> S (IN REF. 2).
 SQ SEQUENCE 353 AA; 40303 MW; 1BF757D865057621 CRC64;
 Query Match 3.3%; Score 7; DB 1; Length 353;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 169 IGLSLIL 175
 DB 202 IGLSLIL 208
 RESULT 20
 OCD_AGRTS STANDARD; PRT; 354 AA.
 ID OCD_AGRTS
 AC P09773;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Ornithine cyclodiaminase (EC 4.3.1.12) (OCD).
 GN ABCB OR OCD OR ATU6016 OR AGR_PTI_54.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OG Plasmid pTiC58.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 OC NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88185308; PubMed=3281832;
 RA Sans N., Schindler U., Schroeder J.;
 RT "Ornithine cyclodiaminase from Ti plasmid C58: DNA sequence, enzyme
 RT properties and regulation of activity by arginine."
 RL Eur. J. Biochem. 173:123-130(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94321320; PubMed=8045881;
 RA Zanker H., Burz G., Langridge U., Langridge P., Kreusch D.,
 RA Schroeder J.;
 RT "Octopine and nopaline oxidases from Ti plasmids of Agrobacterium
 RT tumefaciens: molecular analysis, relationship, and functional
 RT characterization.";
 RL J. Bacteriol. 176:4511-4517(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.B., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eissen J.A., Karp P.D., Bovee D. Sr.,
 RA Chapman P., Clendinning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayavin T., Levy R., Li M.-J., Mclelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,

```
RA Chumley P., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkie G., Gattung S., Miller N., Blanchard M., Mullin L.,
RA Quorllo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmel K., Gordon J., Vaudin M., Tartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RA "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
CC -!- CATALYTIC ACTIVITY: L-ornithine = L-proline + NH(3).
CC -!- COFACTOR: NAD.
CC -!- ENZYME REGULATION: Activity is subject to substrate inhibition
CC and is regulated by L-arginine.
CC -!- PATHWAY: Conversion of nopaline to proline; last step.
CC -!- SIMILARITY: Belongs to the ornithine cyclodeaminase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X07435; CAA30316.1; -
DR EMBL; Z30316; CAA82966.1; -
DR EMBL; AE009420; AAL46252.1; ALT_INIT.
DR EMBL; AE007928; AAK90974.1; ALT_INIT.
DR PIR; S00402; DUAGO.
DR PIR; S55582; S55582.
DR InterPro; IPR003462; ODC_Mu_crystall.
DR Pfam; PF02423; ODC_Mu_crystall; 1.
KW Lyase; NAD; Plasmid; Complete proteome.
FT CONFLICT 212 212 K->E (IN REF. 1 AND 2).
FT CONFLICT 297 297 L->I (IN REF. 1 AND 2).
SQ SEQUENCE 354 AA; 38983 MW; 8272024E0BFDBA24 CRC64;

Query Match 3.3%; Score 7; DB 1; Length 354;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 179 FGVLSDV 185
Db 96 FGVLSDV 102
|||||

RESULT 21
OCD_AGR74
ID -OCD_AGR74 STANDARD; PRT; 356 AA.
AC Q59701;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ornithine cyclodeaminase (EC 4.3.1.12) (OCD).
GN ARCB OR OCD.
OS Agrobacterium tumefaciens (strain Ach5).
OG Plasmid pRIAch5.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176298;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89123162; PubMed=2644238;
RA Schindler U., Sans N., Schroeder J.;
RT "Ornithine cyclodeaminase from octopine Ti plasmid Ach5:
```

```
RT identification, DNA sequence, enzyme properties, and comparison with
RT gene and enzyme from nopaline Ti plasmid C58."
RL J. Bacteriol. 171:847-854(1989).
RN [2]
RP SEQUENCE FROM N.A.
RA Schroeder J., von Lintig J., Zanker H.;
RT "Catabolism of the guanidino compounds nopaline, octopine, and
RT L-arginine in Agrobacterium tumefaciens: enzymes, genes, and
RT regulation."
RL (in) De Deyn P.P., Marescau B., Stalon V., Qureshi I.A. (eds.);
RL Guanidino compounds in biology and medicine, pp.19-28, J. Libbey,
RL London (1992).
CC -!- CATALYTIC ACTIVITY: L-ornithine = L-proline + NH(3).
CC -!- COFACTOR: NAD.
CC -!- ENZYME REGULATION: Activity is subject to substrate inhibition and
CC is regulated by L-arginine.
CC -!- PATHWAY: Conversion of nopaline to proline; last step.
CC -!- SIMILARITY: Belongs to the ornithine cyclodeaminase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M24146; AAA50518.1; -
DR EMBL; Z30328; CAA82989.1; -
DR PIR; A32049; A32049.
DR InterPro; IPR003462; ODC_Mu_crystall.
DR Pfam; PF02423; ODC_Mu_crystall; 1.
KW Lyase; NAD; Plasmid.
SQ SEQUENCE 356 AA; 39166 MW; 5B3D3765E4800395 CRC64;

Query Match 3.3%; Score 7; DB 1; Length 356;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 179 FGVLSDV 185
Db 96 FGVLSDV 102
|||||

RESULT 22
SYN_CLOAB
ID -SYN_CLOAB STANDARD; PRT; 430 AA.
AC Q97FJ7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Histidyl-tRNA synthetase (EC 6.1.1.21) (Histidine--tRNA ligase)
DE (HISRS).
GN HIS OR CAC2740.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.I., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum."
RL J. Bacteriol. 183:4823-4838(2001).
CC -!- CATALYTIC ACTIVITY: ATP + L-histidine + tRNA(His) = AMP +
CC diphosphate + L-histidyl-tRNA(His).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
```

CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE007771; AAK80688.1; -.
DR PIR; C97237; C97237.
DR HAMAP; MF 00127; -; 1.
DR InterPro; IPR004154; HGTP anticodon.
DR InterPro; IPR004516; HISS.
DR InterPro; IPR002314; tRNA synth 2b.
DR Pfam; PF03129; HGTP anticodon; 1.
DR Pfam; PF00587; tRNA-synt 2b; 1.
DR TIGRFAMs; TIGR00442; hisS; 1.
DR PROSITE; PS00862; AA_TRNA_LIGASE II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 430 AA; 49246 MW; D7B034BD474297F3 CRC64;

Query Match 3.3%; Score 7; DB 1; Length 430;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 YDKITEE 28
DB 192 YDKITEE 198
|||||

RESULT 23
Y166 CAEEL
ID Y166 CAEEL STANDARD; PRT; 434 AA.
AC Q11082;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable G protein-coupled receptor B0563.6.
GN B0563.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1] _
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Favella A.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Waterston R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Not known. Putative receptor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U28740; AAA68317.3; -.
DR WormPep; B0563.6; CE29551.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7cm1; 1.
DR PRINTS; PR00237; GPCRHHODOPSIN.

DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Hypothetical protein; G-protein coupled receptor; Transmembrane;
FT TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 65 85 POTENTIAL.
FT TRANSMEM 105 125 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT TRANSMEM 208 228 POTENTIAL.
FT TRANSMEM 258 278 POTENTIAL.
FT CARBOHYD 12 12 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 423 429 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 430 430 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 434 AA; 50026 MW; D5936F6D2470C01D CRC64;

Query Match 3.3%; Score 7; DB 1; Length 434;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVALSLE 145
DB 121 VVALSLE 127
|||||

RESULT 24
PNCB_VIBVU
ID PNCB_VIBVU STANDARD; PRT; 437 AA.
AC Q8DA38;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Nicotinate phosphoribosyltransferase (EC 2.4.2.11) (NAPRTase).
GN PNCB OR VV12372.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1] _
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RL "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Nicotinate D-ribonucleotide + diphosphate =
CC nicotinate + 5-phospho-alpha-D-ribose 1-diphosphate.
CC -!- PATHWAY: NAD biosynthesis; nicotinamide to NAMN; second step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the NAPRTase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE016805; AAO10746.1; -.
DR HAMAP; MF 00570; -; 1.
DR InterPro; IPR007229; NAPRTase.
DR InterPro; IPR006406; Nic_Ptrans.
DR Pfam; PF04095; NAPRTase; 1.
DR TIGRFAMs; TIGR01514; NAPRTase; 1.
KW Pyridine nucleotide biosynthesis; Transferase; Glycosyltransferase;
KW Complete proteome.
SQ SEQUENCE 437 AA; 50001 MW; 6E1DB458931754A7 CRC64;

Query Match 3.3%; Score 7; DB 1; Length 437;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 LKQWKRQ 87
Db 422 LKQWKRQ 428

RESULT 25
ORCS5 DROME STANDARD; PRT; 460 AA.
AC Q24169; Q9V398;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Origin recognition complex subunit 5.
GN ORCS5 OR C37833.

OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96099403; PubMed=7502079;
RA Gossen M., Pak D.T.S., Hansen S.K., Acharya J.K., Botchan M.R.;
RT "A Drosophila homolog of the yeast origin recognition complex.";
RL Science 270:1674-1677 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=99403001; PubMed=10471707;
RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R.G., Davis T.,
Doyle C., Galle R.F., George R.A., Harris N.L., Hartzell G., Martin C.,
Harvey D., Hong L., Houston K.A., Hoskins R.A., Johnson G.,
Moshrefi A., Palazolo M., Reese M.G., Spradling A.C., Tsang G.,
Wan K.H., Whitelaw K., Celniker S.E., Rubin G.M.;
RA "An exploration of the sequence of a 2.9-Mb region of the genome of
RT Drosophila melanogaster: the Adh region.";
RL Genetics 153:179-219 (1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
Fowler A., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,
Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Sampson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
CC -!- FUNCTION: Component of the origin recognition complex (ORC) that
binds origins of replication. It has a role in both chromosomal
replication and mating type transcriptional silencing. Binds to
the ARS consensus sequence (ACS) of origins of replication in an
ATP-dependent manner.
CC -!- SUBUNIT: ORC is composed of six subunits.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the ORC5 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U43505; AAC46956.1; -;
DR EMBL; AE003408; AAF44824.1; -;
DR EMBL; AE003641; AAF53340.1; -;
DR FlyBase; FBN0015271; Orc5.
DR GO; GO:0005664; C:nuclear origin of replication recognition c.; IDA.
DR GO; GO:0006270; P:DNA replication initiation; IDA.
DR GO; GO:0006260; P:DNA replication; IMP.
DR GO; GO:0007076; P:mitotic chromosome condensation; IMP.
DR GO; GO:0007052; P:mitotic spindle assembly; IMP.
DR InterPro; IPR003593; AAA_ATPase.
DR SMART; SM00382; AAA; 1.
DR DNA replication; Nuclear protein; ATP-binding.
KW NP_BIND 41
FT ATP (POTENTIAL).
SQ SEQUENCE 460 AA; 52115 MW; FDCE396E1CBF7D2 CRC64;
Query Match 3.3%; Score 7; DB 1; Length 460;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 KELLIAA 9
Db 326 KELLIAA 332

RESULT 26
CHIO HUMAN STANDARD; PRT; 468 AA.
AC P52757;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Beta-chimaerin (Beta-chimerin).
GN CHN2 OR BCH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RX MEDLINE=94230370; PubMed=8175705;
RA Leung T., How B.-E., Manser E., Lim L.;
RT "Cerebellar beta 2-chimaerin, a GTPase-activating protein for p21
ras-related rac is specifically expressed in granule cells and has a
unique N-terminal SH2 domain.";
RL J. Biol. Chem. 269:12888-12892 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;

RX MEDLINE=95339337; PubMed=7614486;
 RA Yuan S., Miller D.W., Barnett G.H., Hahn J.F., Williams B.R.G.;
 RT "Identification and characterization of human beta 2-chimaerin;
 RL association with malignant transformation in astrocytoma.";
 RI Cancer Res. 55:3456-3461(1995).
 RN [3]
 RP SEQUENCE OF 18-192 FROM N.A.
 RA Strong C., Graves T., Yeakum M., Hawkins M.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC !- FUNCTION: GTPase activating protein for p21-rac. Insufficient
 CC expression of beta-2 chimaerin is expected to lead to higher Rac
 CC activity and could therefore play a role in the progression from
 CC low-grade to high-grade tumors.
 CC !- SUBCELLULAR LOCATION: Membrane-associated (Potential).
 CC !- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=Beta-2;
 CC IsoId=P52757-1; Sequence=Displayed;
 CC Name=Beta-1;
 CC IsoId=P52757-2; Sequence=Not described;
 CC !- TISSUE SPECIFICITY: Highest levels in the brain and pancreas. Also
 CC expressed in the heart, placenta, and weakly in the kidney and
 CC liver. Expression is much reduced in the malignant gliomas,
 CC compared to normal brain or low-grade astrocytomas.
 CC !- SIMILARITY: Contains 1 Rho-GAP domain.
 CC !- SIMILARITY: Contains 1 SH2 domain.
 CC !- SIMILARITY: Contains 1 zinc-dependent phorbol-ester and DAG
 CC binding domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC Use by non-profit institutions. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L29126; AAA19191.1; -;
 CC EMBL; U07223; AAA16836.1; -;
 CC EMBL; U28926; AAA86528.1; -;
 CC EMBL; AC004417; AAC06177.1; -;
 CC PIR; A53764; A53764.
 CC HSP; P28867; IPTQ.
 CC Genew; HGNC:1944; CHN2.
 CC MIM; 602857; -;
 CC GO; GO:0005096; F:GTPase activator activity; TAS.
 CC GO; GO:0005070; F:SH3/SH2 adaptor protein activity; TAS.
 CC InterPro; IPR002219; DAG_P2-bind.
 CC InterPro; IPR008936; Rho-GAP.
 CC InterPro; IPR000198; RhoGAP.
 CC Pfam; PF00130; DAG_P2-bind; 1.
 CC Pfam; PF00620; RhoGAP; 1.
 CC Pfam; PF00017; SH2; 1.
 CC PRINTS; PRO0008; DAGPEDOMAIN.
 CC ProDom; PD000093; SH2; 1.
 CC SMART; SM00109; C1; 1.
 CC SMART; SM00324; RhoGAP; 1.
 CC SMART; SM00252; SH2; 1.
 CC PROSITE; PS00479; DAG_P2_BIND_DOM_1; 1.
 CC PROSITE; PS00081; DAG_P2_BIND_DOM_2; 1.
 CC PROSITE; PS0238; RHO-GAP; 1.
 CC PROSITE; PS0001; SH2; 1.
 CC GTPase activation; Phorbol-ester binding; Zinc; SH2 domain;
 KW Alternative splicing.
 FT DOMAIN 59 127
 FT DOMAIN 215 264
 FT DOMAIN 277 468
 FT CONFLICT 1 6
 FT CONFLICT 468 AA; 53923 MW; 632549580B5804C CRC64;
 SQ SEQUENCE 468 AA; 53923 MW; 632549580B5804C CRC64;
 Query Match 3.3%; Score 7; DB 1; Length 468;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 76 IEARGLK 82
 DB 301 IEARGLK 307
 RESULT 27
 ID ARAA_YERPE STANDARD; PRT; 500 AA.
 AC P58540;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE L-arabinose isomerase (EC 5.3.1.4).
 GN ARAA OR YPO2253 OR Y2094.
 OS Versinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CO-92 / Biovar Orientalis;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Feltwell T., Hamlin N., Holroyd S., Jagsis K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.,
 RA Sammonds M., Skellern J., Stevens K., Whitehead S., Barrell B.G.;
 RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
 RL Nature 413:523-527(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KIMS / Biovar Mediaevalis;
 RX MEDLINE=22137863; PubMed=12142430;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
 RA Perry R.D.;
 RT "Genome sequence of Yersinia pestis KIM.";
 RL J. Bacteriol. 184:4601-4611(2002).
 CC !- CATALYTIC ACTIVITY: L-arabinose = L-ribose.
 CC !- PATHWAY: L-arabinose catabolism; first step.
 CC !- SIMILARITY: Belongs to the arabinose isomerase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC Use by non-profit institutions. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AJ141451; CAC91057.1; -;
 CC EMBL; AE013812; AM85657.1; -;
 CC PIR; AE0274; AE0274.
 CC InterPro; IPR003762; Lara_isomerase.
 CC Pfam; PF02610; Arabinose Isome; 1.
 CC ProDom; PD018364; Lara_isomerase; 1.
 CC Isomerase; Arabinose catabolism; Complete proteome.
 KW Isomerase; Arabinose catabolism; Complete proteome.
 SQ SEQUENCE 500 AA; 55598 MW; 4A26808938FAA45 CRC64;
 Query Match 3.3%; Score 7; DB 1; Length 500;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 169 IGLSIL 175
 DB 90 IGLSIL 96


```
RESULT 28
C4DL_DROME STANDARD; PRT; 511 AA.
AC Q9VLT7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable cytochrome P450 4d21 (EC 1.14.-.-) (CYP1VD21).
GN CYP4D21 OR C65730.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=2019606; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balwle R.M., Basay A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murzy D.M., Nelson D.L.,
RA Nelson D.R., Nelson K., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Poilard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RT Science 287:2185-2195(2000).
CC -!- FUNCTION: May be involved in the metabolism of insect hormones and
CC in the breakdown of synthetic insecticides (By similarity).
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum
CC (Potential).
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
-----
CC EMBL; AF003618; AAF52531.1; -.
CC FlyBase; FBgn0031925; Cyp4d21.
CC InterPro; IPR001128; Cytochrome_P450.
CC Pfam; PF00067; P450; 1.
CC PRINTS; PR00385; P450.
CC PROSITE; PS00086; CYTOCHROME_P450; 1.
CC Oxidoreductase; Monooxygenase; Membrane; Heme; Microsome;
KW Endoplasmic reticulum; Hypothetical protein.
FT METAL 456 456 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 511 AA; 58264 MW; DFL30DC0603BDA50 CRC64;

Query Match 3.3%; Score 7; DB 1; Length 511;
Best Local Similarity 100.0%; Pred.No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 KVLAPAF 206
DB 129 KVLAPAF 135

RESULT 29
CG48 SCHPO STANDARD; PRT; 556 AA.
ID CG48 SCHPO
AC P78750;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Hypothetical WD-repeat protein C29A3.06 in chromosome II.
GN SPBC29A3.06.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Beyer K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RT Nature 415:871-880(2002).
RL Nature 415:871-880(2002).
RN [2]
RP SEQUENCE OF 269-556 FROM N.A.
RC STRAIN=PR745;
RX MEDLINE=98162722; PubMed=9501991;
RA Yoshioka S., Kato K., Nakai K., Okayama H., Nojima H.;
RT "Identification of open reading frames in Schizosaccharomyces pombe
RT cDNAs."
RT DNA Res. 4:363-369(1997).
```

```

CC -1- SIMILARITY: Belongs to the WD-repeat CGI-48 family.
CC -1- SIMILARITY: Contains 5 WD repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL022299; CAA18383.1; -.
DR EMBL; DB0908; BAA13761.1; -.
DR PIR; T40077; T40077.
DR GenDB_Sponbe; SPEC29A3.06; -.
DR InterPro; IPR001860; WD40.
DR Pfam; PF00400; WD40; 4.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00678; WD_REPEATS_1; FALSE_NEG.
DR PROSITE; PS00682; WD_REPEATS_2; 1.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Hypothetical protein; Repeat; WD repeat.
FT REPEAT 216..254
FT REPEAT 259..298
FT REPEAT 300..329
FT REPEAT 330..429
FT REPEAT 438..479
FT REPEAT 485..523
FT REPEAT 523..523
SQ SEQUENCE 556 AA; 62715 MW; 2DF02D921F9D5D7D CRC64;
Query Match 3.3%; Score 7; DB 1; Length 556;
Best Local Similarity 100.0%; Pred.No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 180 GVLSVDL 186
DB 350 GVLSVDL 356
-----
RESULT 30
VIRA AGRT5 STANDARD; PRT; 833 AA.
AC P18540; Q52297;
DT 01-NOV-1990 (Rel. 16, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Wide host range virA protein (EC 2.7.3.-) (WHR virA).
GN VIRA OR ATU6166 OR AGR PTI 2.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OG Plasmid pTiC58.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=90301800; PubMed=2194232;
RA Rogowski P.M., Powell B.S., Shirasu K., Lin T.-S., Morel P.,
RA Zyprian E.M., Steck T.R., Kado C.I.;
RT "Molecular characterization of the vir regulon of Agrobacterium
RT tumefaciens: complete nucleotide sequence and gene organization of
RT the 28.63-kbp region cloned as a single unit.";
RL Plasmid 23:85-106(1990).
RN [2]
RN SEQUENCE FROM N.A.
RA Powell G.K.;
RA Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11741193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Bisen J.A., Karp P.D., Boyee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,

```

```

RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nestor E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58";
RL Science 294:2317-2323(2001).
RN [4]
RN SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmieu K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58";
RL Science 294:2323-2328(2001).
CC -1- FUNCTION: ACTIVATES VIRG, BY PHOSPHORYLATING IT, IN THE
CC PRESENCE OF ACETOSYRINGONE OR HYDROXYSYRINGONE.
CC -1- SIMILARITY: Contains 1 histidine kinase domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03320; AAA91590.1; -.
DR EMBL; L48210; AAA79282.1; -.
DR EMBL; A3009435; AAL46402.1; -.
DR EMBL; A3007923; AAK90927.1; -.
DR PIR; AD3248; AD3248.
DR PIR; S06972; S06972.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003651; His_KinA_N.
DR InterPro; IPR005467; His_Kinase.
DR Pfam; PF02518; HATPase_C_1.
DR Pfam; PF00512; HSKA_1.
DR PRINTS; PR00344; BCTRLENSOR.
DR SMART; SM00387; HATPase_C_1.
DR SMART; SM00388; HSKA_1.
DR PROSITE; PS0109; HIS_KIN; 1.
KW Sensory transduction; Transmembrane; Kinase; Phosphorylation; Plasmid;
KW Crown gall tumor; Transmembrane; Complete proteome.
FT TRANSMEM 19..37
FT DOMAIN 475..698
FT MOD_RES 478..478
FT CONFLICT 68..68
FT CONFLICT 322..322
FT CONFLICT 455..455
FT CONFLICT 455..455
SQ SEQUENCE 833 AA; 91322 MW; B514DAF85BDF32B5 CRC64;
Query Match 3.3%; Score 7; DB 1; Length 833;
Best Local Similarity 100.0%; Pred.No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 141 ALSLEIS 147
DB 181 ALSLEIS 187
-----
RESULT 31
SYA BARBA
ID SYA BARBA STANDARD; PRT; 886 AA.
AC P70865;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)

```

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA ligase) (AlaRS).
GN ALAS.
OS Bartonella bacilliformis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bartonellaceae; Bartonella.
OX NCBI_TaxID=774;
RN [1]
RP SEQUENCE FROM N.A.
RA Upeslatis E., Ihler G.M.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +
CC diphosphate + L-alanyl-tRNA(Ala).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U68242; AAB09037.1; -;
DR HAMAP; MF 00036; -; 1.
DR InterPro; IPR003156; DHHA1.
DR InterPro; IPR002318; tRNA-synt_2c.
DR InterPro; IPR006193; tRNA-synt_Ala.
DR Pfam; PF02272; DHHA1; 1.
DR Pfam; PF01411; tRNA-synt_2c; 1.
DR TIGRfams; TIGR00344; alaS; 1.
DR PROSITE; PSS0860; AA tRNA LIGASE II ALA; 1.
KW Aminoacyl-tRNA synthetase; Protein Biosynthesis; Ligase; ATP-binding.
SQ SEQUENCE 886 AA; 97907 MW; 845AADB4611BCA42 CRC64;

Query Match 3.3%; Score 7; DB 1; Length 886;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 AIDDAIA 38
Db 632 AIDDAIA 638
|||||

RESULT 32
SEC7 YEAST STANDARD; PRT; 2009 AA.
AC P11075; Q03960; Q04139;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein transport protein SEC7.
GN SEC7 OR YDR170C OR YD9395.01C OR YD9489.05C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=88298841; PubMed=3042778;
RX Achstetter T., Franzosoff A., Field C., Schekman R.;
RT "SEC7 encodes an unusual, high molecular weight protein required for
RT membrane traffic from the yeast Golgi apparatus.";
RL J. Biol. Chem. 263:11711-11717(1988).
RN [2]
RP SEQUENCE OF 1-542 FROM N.A.
RC STRAIN=S288c / AB972;
RA Murphy L., Harris D.E., Barrell B.G., Rajandream M.A.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 204-2009 FROM N.A.
RC STRAIN=S288c / AB972;

RA Oliver K., Harris D., Barrell B.G., Rajandream M.A.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=91093340; PubMed=1986005;
RA Franzosoff A., Redding K., Crosby J., Fuller R.S., Schekman R.;
RT "Localization of components involved in protein transport and
RT processing through the yeast Golgi apparatus.";
RL J. Cell Biol. 112:27-37(1991).
CC -!- FUNCTION: May play a role in vesicular budding and traffic between
CC compartments of the Golgi apparatus.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; associated with the peripheral
CC Golgi membrane.
CC -!- DOMAIN: The highly charged acidic domain may serve a structural
CC role to interact with lipids or proteins on the cytoplasmic
CC surface of the Golgi apparatus.
CC -!- PTM: Phosphorylated.
CC -!- SIMILARITY: Contains 1 SEC7 domain.
CC -!- SIMILARITY: TO YEAST YEL022W.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J03918; AAB04031.1; -;
DR EMBL; 246727; CAA86696.1; -;
DR EMBL; 247813; CAA87801.1; -;
DR PIR; S49764; S49764.
DR HSSP; Q99418; IPSV.
DR Germonline; 140661; -;
DR SGD; S0002577; SEC7.
DR GO; GO:0005829; C:cytosol; IDA.
DR GO; GO:0005798; C:Golgi vesicle; IDA.
DR GO; GO:0005770; C:late endosome; IDA.
DR GO; GO:0005086; F:ARF guanyl-nucleotide exchange factor activity; IDA.
DR GO; GO:0006888; F:ER to Golgi transport; IMP.
DR GO; GO:0006891; P:intra-Golgi transport; IMP.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR00904; SEC7.
DR Pfam; PF01369; Sec7; 1.
DR SMART; SM00222; Sec7; 1.
DR PROSITE; PSS0190; SEC7; 1.
KW Transport; Protein transport; Golgi stack; Phosphorylation.
FT DOMAIN 89 213 ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 824 1010 SEC7.
FT CONFLICT 188 188 A -> S (IN REF. 1).
FT CONFLICT 399 400 FV -> LL (IN REF. 1).
FT CONFLICT 402 402 S -> C (IN REF. 1).
FT CONFLICT 1031 1034 QQSA -> PAIC (IN REF. 1).
FT CONFLICT 1036 1037 NF -> QL (IN REF. 1).
SQ SEQUENCE 2009 AA; 226885 MW; 02B2D370DD2E4661 CRC64;

Query Match 3.3%; Score 7; DB 1; Length 2009;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 177 PIFGVLS 183
Db 1588 PIFGVLS 1594
|||||

RESULT 33
LHBI_ECTHL STANDARD; PRT; 40 AA.
ID LHBI_ECTHL
AC P80106;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Light-harvesting protein B800/830/1020, beta-1 chain (EHS-beta-1)

DE (Antenna pigment protein, beta-1 chain) (Fragment).
 OS Ectothiorhodospira halochloris.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales;
 CC Ectothiorhodospiraceae; Halorhodospira.
 CC NCBI_TaxID=1052;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=DSM 1059;
 RX MEDLINE=92249336; PubMed=1577009;
 RA Wagner-Huber R., Brunisholz R.A., Bissig I., Frank G., Suter F.,
 RA Zuber H.;
 RT "The primary structure of the antenna polypeptides of
 RT Ectothiorhodospira halochloris and Ectothiorhodospira halophila. Four
 RT core-type antenna polypeptides in E. halochloris and E. halophila.";
 RL Eur. J. Biochem. 205:917-925(1992).
 CC -!- FUNCTION: ANTENNA COMPLEXES ARE LIGHT-HARVESTING SYSTEMS, WHICH
 CC TRANSFER THE EXCITATION ENERGY TO THE REACTION CENTERS.
 CC -!- SUBUNIT: THE CORE COMPLEX IS FORMED BY DIFFERENT ALPHA AND BETA
 CC CHAINS. BINDING BACTERIOCHLOROPHYLL MOLECULES, AND ARRANGED
 CC MOST PROBABLY IN TETRAMERIC STRUCTURES DISPOSED AROUND THE
 CC REACTION CENTER. THE NON-PIGMENTED GAMMA CHAINS MAY CONSTITUTE
 CC ADDITIONAL COMPONENTS.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.
 CC PIR; S23286; S23286.
 DR HSSP; P02951; 1DX7.
 DR InterPro; IPR000066; Antenna a/b.
 DR InterPro; IPR002362; Antenna_beta.
 DR Prosite; PS00969; ANTENNA COMP BETA; PARTIAL.
 KW Antenna complex; Light-harvesting polypeptide; Transmembrane;
 KW Magnesium; Bacteriochlorophyll; Inner membrane.
 FT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 21 >40
 FT METAL 19 19
 FT METAL 37 37
 FT METAL 40 40
 FT NON TER 40 40
 SQ SEQUENCE 40 AA; 4575 MW; 53536682D22B5D7C CRC64;
 Query Match 2.8%; Score 6; DB 1; Length 40;
 Best Local Similarity 100.0%; Pred.No. 48;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 AVAFVA 14
 DB 31 AVAFVA 36
 RESULT 34
 LH1_RHOTE
 ID LH1_RHOTE STANDARD; PRT; 48 AA.
 AC P80590;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Light-harvesting polypeptide B-885, beta-1 chain (LH-1) (Antenna
 DE pigment polypeptide, beta-1 chain).
 OS Rhodocyclus tenuis (Rhodospirillum tenue).
 CC Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales;
 CC Rhodocyclaceae; Rhodocyclus.
 CC NCBI_TaxID=1066;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=DSM 109;
 RX MEDLINE=96283832; PubMed=8681949;
 RA Hu Q., Brunisholz R.A., Frank G., Zuber H.;
 RT "The antenna complexes of the purple non-sulfur photosynthetic
 RT bacterium Rhodocyclus tenuis. Structural and spectral
 RT characterization.";
 RL Eur. J. Biochem. 238:381-390(1996).
 CC -!- FUNCTION: ANTENNA COMPLEXES ARE LIGHT-HARVESTING SYSTEMS, WHICH
 CC TRANSFER THE EXCITATION ENERGY TO THE REACTION CENTERS.
 CC -!- SUBUNIT: THE CORE COMPLEX IS FORMED BY DIFFERENT ALPHA AND BETA
 CC CHAINS. BINDING BACTERIOCHLOROPHYLL MOLECULES, AND ARRANGED
 CC MOST PROBABLY IN TETRAMERIC STRUCTURES DISPOSED AROUND THE
 CC REACTION CENTER. THE NON-PIGMENTED GAMMA CHAINS MAY CONSTITUTE
 CC ADDITIONAL COMPONENTS.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.
 CC PIR; S68886; S68886.
 DR HSSP; P02951; 1DX7.
 DR InterPro; IPR000066; Antenna a/b.
 DR InterPro; IPR002362; Antenna_beta.
 DR Pfam; PF00556; LHC; 1.
 DR PRINTS; PR00674; LIGHTHARVSTB.
 DR PROSITE; PS00969; ANTENNA COMP BETA; FALSE NEG.
 KW Antenna complex; Light-harvesting polypeptide; Transmembrane;
 KW Magnesium; Bacteriochlorophyll; Inner membrane.
 FT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 21 43
 FT METAL 44 48
 FT METAL 37 37
 CC -!- SUBUNIT: THE CORE COMPLEX IS FORMED BY DIFFERENT ALPHA AND BETA

CC CHAINS. BINDING BACTERIOCHLOROPHYLL MOLECULES, AND ARRANGED
 CC MOST PROBABLY IN TETRAMERIC STRUCTURES DISPOSED AROUND THE
 CC REACTION CENTER. THE NON-PIGMENTED GAMMA CHAINS MAY CONSTITUTE
 CC ADDITIONAL COMPONENTS.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.
 CC PIR; S68885; S68885.
 DR HSSP; P02951; 1DX7.
 DR InterPro; IPR000066; Antenna a/b.
 DR InterPro; IPR002362; Antenna_beta.
 DR Pfam; PF00556; LHC; 1.
 DR PRINTS; PR00674; LIGHTHARVSTB.
 DR PROSITE; PS00969; ANTENNA COMP BETA; FALSE NEG.
 KW Antenna complex; Light-harvesting polypeptide; Transmembrane;
 KW Magnesium; Bacteriochlorophyll; Inner membrane.
 FT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 21 43
 FT METAL 37 37
 FT METAL 44 48
 FT METAL 37 37
 SQ SEQUENCE 48 AA; 5344 MW; 8CB2F1C3736F086 CRC64;
 Query Match 2.8%; Score 6; DB 1; Length 48;
 Best Local Similarity 100.0%; Pred.No. 56;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 10 VAFVAV 15
 DB 26 VAFVAV 31
 RESULT 35
 LH2_RHOTE
 ID LH2_RHOTE STANDARD; PRT; 48 AA.
 AC P80591;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Light-harvesting polypeptide B-885, beta-2 chain (LH-1) (Antenna
 DE pigment polypeptide, beta-2 chain).
 OS Rhodocyclus tenuis (Rhodospirillum tenue).
 CC Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales;
 CC Rhodocyclaceae; Rhodocyclus.
 CC NCBI_TaxID=1066;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=DSM 109;
 RX MEDLINE=96283832; PubMed=8681949;
 RA Hu Q., Brunisholz R.A., Frank G., Zuber H.;
 RT "The antenna complexes of the purple non-sulfur photosynthetic
 RT bacterium Rhodocyclus tenuis. Structural and spectral
 RT characterization.";
 RL Eur. J. Biochem. 238:381-390(1996).
 CC -!- FUNCTION: ANTENNA COMPLEXES ARE LIGHT-HARVESTING SYSTEMS, WHICH
 CC TRANSFER THE EXCITATION ENERGY TO THE REACTION CENTERS.
 CC -!- SUBUNIT: THE CORE COMPLEX IS FORMED BY DIFFERENT ALPHA AND BETA
 CC CHAINS. BINDING BACTERIOCHLOROPHYLL MOLECULES, AND ARRANGED
 CC MOST PROBABLY IN TETRAMERIC STRUCTURES DISPOSED AROUND THE
 CC REACTION CENTER. THE NON-PIGMENTED GAMMA CHAINS MAY CONSTITUTE
 CC ADDITIONAL COMPONENTS.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.
 CC PIR; S68886; S68886.
 DR HSSP; P02951; 1DX7.
 DR InterPro; IPR000066; Antenna a/b.
 DR InterPro; IPR002362; Antenna_beta.
 DR Pfam; PF00556; LHC; 1.
 DR PRINTS; PR00674; LIGHTHARVSTB.
 DR PROSITE; PS00969; ANTENNA COMP BETA; FALSE NEG.
 KW Antenna complex; Light-harvesting polypeptide; Transmembrane;
 KW Magnesium; Bacteriochlorophyll; Inner membrane.
 FT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 21 43
 FT METAL 44 48
 FT METAL 37 37
 CC -!- SUBUNIT: THE CORE COMPLEX IS FORMED BY DIFFERENT ALPHA AND BETA

```
FT SEQUENCE 48 AA; 5358 MW; 8CB2F1C3736F1825 CRC64;
SQ SEQUENCE 48 AA; 5358 MW; 8CB2F1C3736F1825 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 VAFVAV 15
DB 26 VAFVAV 31

RESULT 36
A95E DROME STANDARD; PRT; 52 AA.
AC P16548; Q9VCB0;
DT 01-AUG-1990 (Rel. 15, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Accessory gland-specific peptide 95EF precursor (Male accessory gland
DE secretory protein 316).
GN ACP95EF OR MST95E OR MSP16 OR CG17924.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S;
RX MEDLINE=90228604; PubMed=2109712;
RA Dibenedetto A.J., Harada H.A., Wolfner M.F.;
RT "Structure, cell-specific expression, and mating-induced regulation
RT of a Drosophila melanogaster male accessory gland gene.";
RL Dev. Biol. 139:134-148 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Gallie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yeaman M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan H.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flesler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei X.-H., Ibegwam C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattet B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M., Paillet M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
```

```
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of Drosophila melanogaster.";
RL Science 287:2183-2195 (2000).
CC -!- FUNCTION: This protein may be a precursor of secreted proteins
CC and peptide hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Main cells of the accessory glands of males.
CC -!- DEVELOPMENTAL STAGE: In very late pupae and in adults.
CC -!- INDUCTION: By mating.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M32022; AAD13400.1; -.
DR EMBL; AE003746; AAP56263.2; -.
DR PIR; A37354; A37354.
DR FlyBase; FBgn0002863; Acp95EF.
KW Behavior; Hormone; Signal.
FT SIGNAL 1 22
FT CHAIN 23 52
FT CONFLICT 20 20
FT SEQUENCE 52 AA; 5405 MW; A44008BD6057FF2B CRC64;
SQ SEQUENCE 52 AA; 5405 MW; A44008BD6057FF2B CRC64;

Query Match 2.8%; Score 6; DB 1; Length 52;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVALSL 144
DB 13 VVALSL 18

RESULT 37
ACP_BACHD STANDARD; PRT; 77 AA.
AC Q9K404;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acyl carrier protein (ACP).
GN ACPP OR ACPA OR BH2490.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirana C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331 (2000).
CC -!- FUNCTION: Carrier of the growing fatty acid chain in fatty
CC acid biosynthesis (By similarity).
CC -!- PATHWAY: De novo fatty acid biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- PTM: 4'-phosphopantetheine is transferred from CoA to a specific
CC serine of apo-ACP by acpS. This modification is essential for
CC activity because fatty acids are bound in thioester linkage to the
CC sulphydryl of the prosthetic group (By similarity).
CC -!- SIMILARITY: Contains 1 acyl carrier domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF001515; BAB06209.1; -
 DR PIR; B83961; B83961.
 DR HSSP; P80643; 1HY8.
 DR HAMAP; MF 01217; -; 1.
 DR InterPro; IPR003231; Acyl carrier.
 DR InterPro; IPR006163; Pp_bind.
 DR InterPro; IPR006162; Ppanthe S.
 DR Pfam; PF00550; pp-binding; 1.
 DR ProDom; PD000887; Acyl carrier; 1.
 DR TIGRFAMs; TIGR00517; acyl carrier; 1.
 DR PROSITE; PS00075; ACP DOMAIN; 1.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
 KW Lipid synthesis; Fatty acid biosynthesis; Phosphopantetheine;
 KW Complete proteome.
 FT BINDING 37 37 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 SQ SEQUENCE 77 AA; 8558 MW; 3575CDFA45BA15CA CRC64;

Query Match 2.8%; Score 6; DB 1; Length 77;
 Best Local Similarity 100.0%; Pred. No. 83;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 LEISDE 149
 DB 53 LEISDE 58

RESULT 38 ATPH_AWTPSP

ID ATPH_AWTPSP STANDARD; PRT; 82 AA.
 AC Q02851;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE ATP synthase C chain (EC 3.6.3.14) (Lipid-binding protein) (Subunit III).
 DE ATP synthase C chain (EC 3.6.3.14) (Lipid-binding protein) (Subunit III).
 GN ATPH.
 OS Anthamion sp.
 OS Chloroplast.
 OC Eukaryota; Rhodophyta; Floridophyceae; Ceramiales; Ceramiaceae;
 OC Anthamion.
 OC NCBI_TaxID=2767;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LB 95.79;
 RX MEDLINE=93021132; PubMed=1404401;
 RA Kostrewa M., Zetsche K.;
 RT "Large ATP synthase operon of the red alga Anthamion sp. resembles the corresponding operon in cyanobacteria."
 RL J. Mol. Biol. 227:961-970(1992).
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate + H(+) (Out).
 CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic core - and CF(0) - the membrane proton channel. CF(1) has five subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0) has three main subunits: a, b and c.
 CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
 CC -!- MISCELLANEOUS: Dicyclohexylcarbodiimide (DCCD) inhibits ATPase.
 CC -!- SIMILARITY: Belongs to the ATPase C chain family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X63382; CAA44980.1; -

DR PIR; S26958; S26958.
 DR InterPro; IPR005953; ATP synth C.
 DR InterPro; IPR002379; ATPase_Csub.
 DR InterPro; IPR000454; Eub_ATPase_Csub.
 DR Pfam; PF00137; ATP-synt C; 1.
 DR PRINTS; PR00124; ATPASEC.
 DR TIGRFAMs; TIGR01260; ATP synt c; 1.
 DR PROSITE; PS00605; ATPase_C; 1.
 KW Hydrogen ion transport; Lipid-binding; Chloroplast; CF(0);
 KW Transmembrane.
 FT TRANSMEM 7 33 BY SIMILARITY.
 FT BINDING 51 77 BY SIMILARITY.
 FT BINDING 61 61 DICYCLOHEXYLCARBODIIMIDE (BY SIMILARITY).
 SQ SEQUENCE 82 AA; 8100 MW; 3F22042A52DE9622 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 82;
 Best Local Similarity 100.0%; Pred. No. 87;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVALSL 144
 DB 69 VVALSL 74

RESULT 39 ATPH_CVACA

ID ATPH_CVACA STANDARD; PRT; 82 AA.
 AC Q9TM30;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE ATP synthase C chain (EC 3.6.3.14) (Lipid-binding protein) (Subunit III).
 DE ATP synthase C chain (EC 3.6.3.14) (Lipid-binding protein) (Subunit III).
 GN ATPH.
 OS Cyanidium caldarium.
 OS Chloroplast.
 OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
 OC Cyanidium.
 OC NCBI_TaxID=2771;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RRK-1;
 RX MEDLINE=20496959; PubMed=11040290;
 RA Gloeckner G., Rosenthal A., Valentin K.-U.;
 RT "The structure and gene repertoire of an ancient red algal plastid genome."
 RL J. Mol. Evol. 51:382-390(2000).
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate + H(+) (Out).
 CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic core - and CF(0) - the membrane proton channel. CF(1) has five subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0) has three main subunits: a, b and c.
 CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
 CC -!- MISCELLANEOUS: Dicyclohexylcarbodiimide (DCCD) inhibits ATPase.
 CC -!- SIMILARITY: Belongs to the ATPase C chain family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF022186; AAF13009.1; -
 DR InterPro; IPR005953; ATP synth C.
 DR InterPro; IPR002379; ATPase_Csub.
 DR InterPro; IPR000454; Eub_ATPase_Csub.
 DR Pfam; PF00137; ATP-synt C; 1.
 DR PRINTS; PR00124; ATPASEC.
 DR TIGRFAMs; TIGR01260; ATP synt c; 1.
 DR PROSITE; PS00605; ATPase_C; 1.

KW Hydrogen ion transport; Lipid-binding; Chloroplast; CF(0);
KW Transmembrane. 7 33 POTENTIAL.
FT TRANSMEM 51 77 POTENTIAL.
FT BINDING 61 61 DICYCLOHEXYLCARBODIIMIDE (BY SIMILARITY).
SQ SEQUENCE 82 AA; 8149 MW; 908E74D1BB0137CF CRC64;

Query Match 2.8%; Score 6; DB 1; Length 82;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVALSL 144
DB 69 VVALSL 74

RESULT 40
ATPH PORPU STANDARD; PRT; 82 AA.
AC P51246; (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP synthase C chain (EC 3.6.3.14) (Lipid-binding protein) (Subunit III).
DE III).
GN ATPH.
OS Porphyra purpurea.
OG Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Porphyra.
OX NCBI_TaxID=2787;
RN SEQUENCE FROM N.A.
RC STRAIN=Avonport;
RA Reith M.E., Munnolland J.;
RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast genome.";
RL Plant Mol. Biol. Rep. 13:333-335(1995).
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate + H(+) (Out).
CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic core and CF(0) - the membrane proton channel. CF(1) has five subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0) has three main subunits: a, b and c.
CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -!- MISCELLANEOUS: Dicyclohexylcarbodiimide (DCDD) inhibits ATPase.
CC -!- SIMILARITY: Belongs to the ATPase C chain family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U38804; AAC08132.1; -
CC F1R; S73167; S73167.
CC InterPro; IPR005953; ATP_synth_C.
CC InterPro; IPR002379; ATPase_Csub.
CC Pfam; PF00137; ATP-synt_C; 1.
CC PRINTS; PR00124; ATPASEC.
CC TIGRFAMs; TIGR01260; ATP_synt_c; 1.
CC PROSITE; PS00605; ATPASE_C; 1.
CC Hydrogen ion transport; Lipid-binding; Chloroplast; CF(0);
KW Transmembrane.
FT TRANSMEM 7 33 POTENTIAL.
FT TRANSMEM 51 77 POTENTIAL.
FT BINDING 61 61 DICYCLOHEXYLCARBODIIMIDE (BY SIMILARITY).
SQ SEQUENCE 82 AA; 8084 MW; 295BD8FA52C62D81 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 82;
Best Local Similarity 100.0%; Pred. No. 87;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 139 VVALSL 144
DB 69 VVALSL 74

RESULT 41
ATPH GALSU STANDARD; PRT; 83 AA.
AC P35013;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ATP synthase C chain (EC 3.6.3.14) (Lipid-binding protein) (Subunit III).
DE III).
GN ATPH.
OS Galdieria sulphuraria (Red alga).
OG Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
OC Galdieria;
OX NCBI_TaxID=130081;
RN SEQUENCE FROM N.A.
RC STRAIN=14-1-1 / Isolate 107.79/Goettingen;
RX MEDLINE=94033298; PubMed=8219057;
RA Kozrzewa M., Zetsche K.;
RT "Organization of plastid-encoded ATPase genes and flanking regions including homologues of infB and tsf in the thermophilic red alga Galdieria sulphuraria.";
RL Plant Mol. Biol. 23:67-76(1993).
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate + H(+) (Out).
CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic core and CF(0) - the membrane proton channel. CF(1) has five subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0) has three main subunits: a, b and c.
CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -!- MISCELLANEOUS: Dicyclohexylcarbodiimide (DCDD) inhibits ATPase.
CC -!- SIMILARITY: Belongs to the ATPase C chain family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X67814; CAA48021.1; -
CC InterPro; IPR005953; ATP_synth_C.
CC InterPro; IPR002379; ATPase_Csub.
CC Pfam; PF00137; ATP-synt_C; 1.
CC PRINTS; PR00124; ATPASEC.
CC TIGRFAMs; TIGR01260; ATP_synt_c; 1.
CC PROSITE; PS00605; ATPASE_C; 1.
CC Hydrogen ion transport; Lipid-binding; Chloroplast; CF(0);
KW Transmembrane.
FT TRANSMEM 7 33 POTENTIAL.
FT TRANSMEM 51 77 POTENTIAL.
FT BINDING 61 61 DICYCLOHEXYLCARBODIIMIDE (BY SIMILARITY).
SQ SEQUENCE 83 AA; 8282 MW; CEC80894FB998E76 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 83;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVALSL 144
DB 69 VVALSL 74

RESULT 42
ATPH PAVLU STANDARD; PRT; 83 AA.
AC P28530;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP synthase C chain (EC 3.6.3.14) (Lipid-binding protein) (Subunit III).
DE I111.
GN ATPH.
OS Pavlova lutherii (Monochrysis lutheri).
OG Chloroplast.
OC Eukaryota; Haptophyceae; Pavlova.
OX NCBI_TaxID=2832;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92316212; PubMed=1618309;
RA Scaramuzzi C.D., Stokes H.W., Hiller R.G.;
RT "Characterisation of a chloroplast-encoded secY homologue and atph from a chromophytic alga. Evidence for a novel chloroplast genome organisation.";
RT FEBS Lett. 304:119-123(1992).
RL FTS Lett. 304:119-123(1992).
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate + H(+) (Out).
CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic core - and CF(0) - the membrane proton channel. CF(1) has five subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0) has three main subunits: a, b and c.
CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -!- MISCELLANEOUS: Dicyclohexylcarbodiimide (DCCD) inhibits ATPase.
CC -!- SIMILARITY: Belongs to the ATPase C chain family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
DR EMBL; X64731; CAB45997.1; -.
DR PIR; S23424; S23424.
DR InterPro; IPR005953; ATP_synth_C.
DR InterPro; IPR002379; ATPase_Csub.
DR Pfam; PF00137; ATP-synt_C; 1.
DR PRINTS; PR00124; ATPASEC.
DR PROSITE; TIGR01260; ATP_synth_C; 1.
DR PROSITE; PS00605; ATPASE_C; 1.
KW Hydrogen ion transport; Lipid-binding; Chloroplast; CF(0); Transmembrane.
FT BINDING 61 61 DICYCLOHEXYLCARBODIIMIDE (BY SIMILARITY).
SQ SEQUENCE 83 AA; 8180 MW; B41F6C0ABA086493 CRC64;
Query Match 2.8%; Score 6; DB 1; Length 83;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 139 VVALSL 144
DB 69 VVALSL 74
RESULT 43
ID YK45_PSEAE STANDARD; PRT; 86 AA.
AC Q91270;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0161 protein PA2045.
GN PA2045.
OS Pseudomonas aeruginosa.

Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P., Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M., Garber R.L., Goltz L., Tolentino B., Westbrook-Wadman S., Yuan Y., Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M., Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.";
RT Nature 406:959-964(2000).
RL Nature 406:959-964(2000).
CC -!- SIMILARITY: Belongs to the UPF0161 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
DR EMBL; AB004631; AAG05433.1; -.
DR PIR; D83389; D83389.
DR HAMAP; MF_00386; -; 1.
DR InterPro; IPR002696; DUF37.
DR Pfam; PF01809; DUF37; 1.
DR ProDom; PD004225; DUF37; 1.
DR TIGRfam; TIGR00278; TIGR00278; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 86 AA; 9824 MW; 5E228B8ED7F5A829 CRC64;
Query Match 2.8%; Score 6; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 MKFLLI 7
DB 1 MKFLLI 6
RESULT 44
ID PTHP_LACCA STANDARD; PRT; 88 AA.
AC Q9KJW3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphocarrier protein Hpr (Histidine-containing protein).
GN PTHP
OS Lactobacillus casei.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1582;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 393 / DSM 20011 / NCDO 161;
RX MEDLINE=20305547; PubMed=10844647;
RA Viana R., Monedero V., Dossennet V., Vadeboncoeur C., Perez-Martinez G., Deutscher J.;
RT "Enzyme I and Hpr from Lactobacillus casei: their role in sugar transport, carbon catabolite repression and inducer exclusion.";
RL Mol. Microbiol. 36:570-584(2000).
CC -!- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOENOLPYRUVATE-DEPENDENT SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYDRATE ACTIVE TRANSPORT SYSTEM. THE PHOSPHORYL GROUP CARRIER PROTEIN HPR BY (PEP) IS TRANSFERRED TO THE PHOSPHORYL CARRIER PROTEIN HPR BY ENZYME I. PHOSPHO-HPR THEN TRANSFERS IT TO THE PERMEASE (ENZYMES II/III). HPR IS COMMON TO ALL PTS (BY SIMILARITY).
CC

CC -1- ENZYME REGULATION: PHOSPHORYLATION ON SER-46 INHIBITS THE
CC PHOSPHORYL TRANSFER FROM ENZYME I TO HPR (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the Hpr family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; AF159589; AAF74346.1; -.
CC HSPSP; P07515; IPTF.
CC InterPro; IPR001020; Hpr_Hisp_S.
CC InterPro; IPR000032; Hpr_protein.
CC InterPro; IPR002114; Hpr_Serp_S.
CC InterPro; IPR005698; PTS_Hpr.
CC Pfam; PF00381; PTS_Hpr; 1.
CC PRINTS; P00107; PHOSPHOCFHR.
CC ProDom; PD002238; Hpr_protein; 1.
CC TIGRFAMs; TIGR01003; PTS_Hpr_family; 1.
CC PROSITE; PS00369; PTS_HPR_HIS; 1.
CC PROSITE; PS00583; PTS_HPR_SER; 1.
CC Phosphotransferase system; Sugar transport; Phosphorylation.
CC MOD_RES 15 15 PHOSPHORYLATION (BY ENZYME I) (BY
CC SIMILARITY).
CC FT MOD_RES 46 46 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 88 AA; 9253 MW; 5723EBE1345F56CA CQC64;
CC
CC Query Match 2.8%; Score 6; DB 1; Length 88;
CC Best Local Similarity 100.0%; Pred. No. 93;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 35 DATAAI 40
CC DB 72 DATAAI 77
CC
CC
CC RESULT 45
CC ESXI MYCBO STANDARD; PRT; 94 AA.
CC ID ESXI MYCBO
CC AC P59802;
CC DT 10-OCT-2003 (Rel. 42, Created)
CC DT 10-OCT-2003 (Rel. 42, Last sequence update)
CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC DE ESAT-6 like protein esxi.
CC GN ESXI OR MB1066C.
CC OS Mycobacterium bovis.
CC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
CC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
CC NCBI_TaxID=1765;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN=AF2122/97;
CC MEDLINE=22703107; PubMed=12788972;
CC Garnier T., Biglmeier K., Camus J.-C., Medina N., Mansoor H.,
CC Pryor M., Duchoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
CC Harris B., Akin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
CC Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
CC "The complete genome sequence of Mycobacterium bovis";
CC Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -1- SIMILARITY: Belongs to the ESAT-6 (esx) family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

DR EMBL; BX248337; CAD93926.1; -.
KW Complete proteome.
SQ SEQUENCE 94 AA; 9818 MW; 8F9701831E93F657 CQC64;
Query Match 2.8%; Score 6; DB 1; Length 94;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 193 SDVLTA 188
DB 33 SDVLTA 38
RESULT 46
ESXI MYCTU STANDARD; PRT; 94 AA.
ID ESXI MYCTU
AC P96364; O08120; O08122; Q9L781;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ESAT-6 like protein esxi (Antigen Mtb9.9B).
GN (RV1037C OR MT1066 OR MTCY10G2.12) AND
GN (RV3619C OR MT3721 OR MTCY15C10.33 OR MTCY07H7B.03).
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1773;
[1]
SEQUENCE FROM N.A.
STRAIN=Erdmann;
MEDLINE=20130298; PubMed=10662800;
Alderson M.R., Bement T., Day C.H., Zhu L., Moles D., Skeiky Y.A.W.,
Coler R., Lewinson D.M., Reed S.G., Dillon D.C.;
"Expression cloning of an immunodominant family of Mycobacterium
tuberculosis antigens using human CD4(+) T cells";
J. Exp. Med. 191:551-560(2000).
[2]
SEQUENCE FROM N.A.
STRAIN=H37RV;
MEDLINE=98295987; PubMed=9634230;
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
Baddock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares R., Squares R.,
Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
"Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence";
Nature 393:537-544(1998).
[3]
SEQUENCE FROM N.A.
STRAIN=CDC 1551 / Oshkosh;
MEDLINE=22206494; PubMed=12218036;
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
Kolonyak J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
"Whole-genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains";
J. Bacteriol. 184:5479-5490(2002).
-1- SIMILARITY: Belongs to the ESAT-6 (esx) family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

```
DR EMBL; AF226277; AAF32406.1; -
DR EMBL; Z92539; CAB06842.1; -
DR EMBL; Z95436; CAB08822.1; -
DR EMBL; AE006989; AAK45317.1; -
DR EMBL; AE007171; AAK48082.1; -
DR PIR; D70560; D70560.
DR TIGR; MT1086; -
DR TIGR; MT3721; -
DR Tuberculist; Rv1037c; -
DR Tuberculist; Rv3619c; -
KW Hypothetical protein; Complete proteome.
FT CONFLICT 20 20 S -> L (IN REF. 1 AND 3; AAK48082).
FT CONFLICT 23 23 S -> L (IN REF. 1 AND 3; AAK48082).
SQ SEQUENCE 94 AA; 9833 MW; 8F971BB03789C57E CRC64;

Query Match 2.8%; Score 6; DB 1; Length 94;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 183 SDVLTJA 188
DB 33 SDVLTJA 38

RESULT 47
NULC_PLEBO STANDARD; PRT; 101 AA.
AC Q00244.
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NAD(P)H-quinone oxidoreductase chain 4L (EC 1.6.5.-) (NAD(P)H
DE dehydrogenase I, chain 4L) (NDH-1, chain 4L).
GN NDHE OR NDH4L.
OS Plectonema boryanum.
OC Bacteria; Cyanobacteria; Oscillatoriales; Plectonema.
OX NCBI_TaxID=1184;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IAM M-101;
RA Takahashi Y., Shonai F., Fujita Y., Kohchi T., Ohyama K.,
RA Matsubara H.;
RT Structure of a co-transcribed gene cluster, ndh1-frxB-ndh6-ndh4L,
RT cloned from the filamentous cyanobacterium Plectonema boryanum.;
RL Plant Cell Physiol. 32:969-981(1991).
CC -!- FUNCTION: NDH-1 shuttles electrons from NAD(P)H, via FMN and iron-
CC sulfur (Fe-S) centers, to quinones in the respiratory chain. The
CC immediate electron acceptor for the enzyme in this species is
CC believed to be plastoquinone. Couples the redox reaction to proton
CC translocation (for every two electrons transferred, four hydrogen
CC ions are translocated across the cytoplasmic membrane), and thus
CC conserves the redox energy in a proton gradient (By similarity).
CC -!- CATALYTIC ACTIVITY: NAD(P)H + plastoquinone = NAD(P)(+) +
CC plastoquinol.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF158101; AAD42645.1; -
CC DR EMBL; AF158101; AAD42645.1; -
KW Hypothetical protein.
SQ SEQUENCE 117 AA; 13135 MW; D8B52577D7DA979B CRC64;

Query Match 2.8%; Score 6; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 AIAAIE 41
DB 69 AIAAIE 74

RESULT 49
GCSE_THEME STANDARD; PRT; 124 AA.
AC Q9WY55;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Glycine cleavage system H protein.
GN GCYH OR TMO212.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109 / ATCC 43589;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";

Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 FLIIAA 9
DB 6 FLIIAA 11

RESULT 48
Y13K_BPT4 STANDARD; PRT; 117 AA.
ID Y13K_BPT4
AC P39504;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical 13.1 kDa protein in pser-alc intergenic region.
GN Y13K OR PSET.3.
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
OX NCBI_TaxID=10685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22514363; PubMed=12626685;
RA Miller E.S., Kutter E., Mosig G., Arisaka F., Kunisawa T., Ruger W.;
RT "Bacteriophage T4 genome.";
RL Microbiol. Mol. Biol. Rev. 67:86-156(2003).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF158101; AAD42645.1; -
CC DR EMBL; AF158101; AAD42645.1; -
KW Hypothetical protein.
SQ SEQUENCE 117 AA; 13135 MW; D8B52577D7DA979B CRC64;
```

RL Nature 399:323-329(1999).
CC -!- FUNCTION: The glycine cleavage system catalyzes the degradation of
CC glycine. The H protein shuttles the methylamine group of glycine
CC from the P protein to the T protein (By similarity).
CC -!- COFACTOR: The H chain contains a covalently-bound lipoyl cofactor
CC (By similarity).
CC -!- SUBUNIT: The glycine cleavage system is composed of four proteins:
CC P, T, L and H (By similarity).
CC -!- SIMILARITY: Belongs to the gcvH family.
CC -!- SIMILARITY: Contains 1 lipoyl-binding domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB001706; AA035304.1; -.
CC PIR; F72403; F72403.
CC HSSP; P16048; LHTP.
CC TIGR; TM0212; -.
CC HAMAP; MF 00272; -; 1.
CC InterPro; IPR002930; GCV_H.
CC InterPro; IPR003016; Lipoyl_BS.
CC Pfam; PF01597; GCV_H; 1.
CC TIGRFAMs; TIGR00527; gcvH; 1.
CC PROSITE; PS00189; LIPOYL; 1.
CC LIPOYL; Complete proteome.
CC BINDING 60 60 LIPOYL (BY SIMILARITY).
CC SEQUENCE 124 AA; 13915 MW; 90B7CDE430A16C9F CRC64;
CC
CC Query March 2.8%; Score 6; DB 1; Length 124;
CC Best Local Similarity 100.0%; Pred. No. 1.2e+02;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 145 EISDEG 150
CC DB 101 EISDEG 106
CC
CC RESULT 50
CC RS6_CHLTF STANDARD; PRT; 131 AA.
CC ID RS6_CHLTF
CC OS CHLTF
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DE 30S ribosomal protein S6.
CC RPSF CR CT2135.
CC OS Chlorobium tepidum.
CC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
CC Chlorobium.
CC NCBI_TaxID=1097;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN=TLS / ATCC 49652 / DSM 12025;
CC MEDLINE=22103685; PubMed=12093901;
CC Eisen J.A., Nelson K.E., Paulsen I.T., Heigelberg J.P., Wu M.,
CC Dodson R.J., Deboy R., Gwin M.L., Nelson W.C., Haft D.H.,
CC Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
CC Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
CC Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
CC Vanathavan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
CC Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
CC "The complete genome sequence of Chlorobium tepidum TLS, a
CC photosynthetic, anaerobic green-sulfur bacterium."
CC Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
CC -!- FUNCTION: Binds together with S18 to 16S ribosomal RNA (By
CC similarity).
CC -!- SIMILARITY: Belongs to the S6P family of ribosomal proteins.
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB012962; AAM73351.1; -.
CC TIGR; CT2135; -.
CC HAMAP; MF 00360; -; 1.
CC InterPro; IPR000529; Ribosomal_S6.
CC Pfam; PF01250; Ribosomal_S6; 1.
CC ProDom; PD003809; Ribosomal_S6; 1.
CC TIGRFAMs; TIGR00166; S6; 1.
CC PROSITE; PS01048; RIBOSOMAL_S6; FALSE NEG.
CC KW Ribosomal protein; rRNA-binding; Complete proteome.
CC SEQUENCE 131 AA; 14741 MW; 2166855AD3F80577 CRC64;
CC
CC Query March 2.8%; Score 6; DB 1; Length 131;
CC Best Local Similarity 100.0%; Pred. No. 1.3e+02;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 34 DDAAIAA 39
CC DB 19 DDAAIAA 24
CC
CC RESULT 51
CC RUVX_HELPJ STANDARD; PRT; 134 AA.
CC ID RUVX_HELPJ
CC OS RUVX_HELPJ
CC DT 10-OCT-2003 (Rel. 42, Created)
CC DT 10-OCT-2003 (Rel. 42, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Putative Holliday junction resolvase (EC 3.1.1.-).
CC GN JHP0317.
CC OS Helicobacter pylori J99 (Campylobacter pylori J99).
CC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
CC Helicobacteraceae; Helicobacter.
CC NCBI_TaxID=85963;
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE=99120557; PubMed=9923682;
CC Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
CC Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
CC Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
CC Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
CC Trust T.J.;
CC "Genomic sequence comparison of two unrelated isolates of the human
CC gastric pathogen Helicobacter pylori."
CC Nature 397:176-180(1999).
CC -!- FUNCTION: Could be a nuclease that resolves Holliday junction
CC intermediates in genetic recombination.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Belongs to the ysgF HUR family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF001468; AAD05908.1; -.
CC PIR; E71945; E71945.
CC HAMAP; MF 00651; -; 1.
CC InterPro; IPR005227; Cons_hypoth250.
CC InterPro; IPR006641; YggFC.
CC Pfam; PF03652; UPF0081; 1.
CC SMART; SM00732; YggFC; 1.
CC Hydrolase; Nuclease; DNA repair; DNA recombination; Complete proteome.
CC

SQ SEQUENCE 134 AA; 15121 MW; C4DF777908C983D2 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 134;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 VDFKGE 70
DB 75 VDFKGE 80

RESULT 52

RUVX HELPY STANDARD; PRT; 134 AA.

AC 025101; (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative Holliday junction resolvase (EC 3.1.-.-).
GN HP0334.

OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=26695 / ATCC 700392;

RX MEDLINE=97394467; PubMed=9252185;

RA Tomb J.-F., White O., Karlavage A.R., Clayton R.A., Sutton G.G., B.A.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glöck A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;

RT "The complete genome sequence of the gastric pathogen Helicobacter
pylori."

RL Nature 388:539-547(1997).

CC -!- FUNCTION: Could be a nuclease that resolves Holliday junction

intermediates in genetic recombination.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).

CC -!- SIMILARITY: Belongs to the ygfR HJR family.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AE000551; AAD07403.1; -

DR PIR; F64561; F64561.

DR TIGR; HP0334; -

DR HAMAP; MF 00651; -

DR InterPro; IPR005227; Cons hypoch250.

DR InterPro; IPR006641; YqgFC.

DR Pfam; PF03652; YqgF081; 1.

DR SMART; SM00732; YqgF081; 1.

DR TIGRFAMs; TIGR00250; YqgF081; 1.

DR TIGRFAMs; TIGR00250; YqgF081; 1.

KW Hydrolase; Nuclease; DNA repair; DNA recombination; Complete proteome.

SQ SEQUENCE 134 AA; 15256 MW; 516D8C4233B932CE CRC64;

Query Match 2.8%; Score 6; DB 1; Length 134;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 VDFKGE 70
DB 75 VDFKGE 80

RESULT 53

JANA DROME

ID JANA DROME STANDARD; PRT; 135 AA.

AC P20348; Q9VAB6;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Sex-regulated protein janus-A.

GN JANA OR CG7933.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Canton-S;

RX MEDLINE=89343970; PubMed=2503707;

RA Yanicostas C., Vincent A., Lepesant J.-A.;

RT "Transcriptional and posttranscriptional regulation contributes to

the sex-regulated expression of two sequence-related genes at the

RT janus locus of Drosophila melanogaster."

RL Mol. Cell. Biol. 9:2526-2535(1989).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkeley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champagne B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brotter P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,

RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarly C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster."

Science 287:2185-2195(2000).

CC -!- TISSUE SPECIFICITY: SOMATIC AND GERM LINE CELLS.

CC -!- DEVELOPMENTAL STAGE: THE NON-SEX-SPECIFIC TRANSCRIPT IS PRESENT AT

ALL STAGES.

CC -!- MISCELLANEOUS: TRANSCRIPTION OF JANA GIVES RISE TO TWO

SEX-SPECIFIC AND TO ONE NON-SEX-SPECIFIC TRANSCRIPTS.

CC -!- SIMILARITY: Belongs to the janus family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M27033; AAC34203.1; --
 DR EMBL; M27033; AAC34202.1; ALT_INIT.
 DR EMBL; AE003772; AAF56997.1; --
 DR PIR; A32317; A32317.
 DR FlyBase; FBgn0001280; jana.
 DR InterPro; IPR007702; Ocnus.
 DR Pfam; PF05005; Ocnus; 1.
 SQ SEQUENCE 135 AA; 15220 MW; 2720237C77F7F3132 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 135;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 EISDEG 150
 DB 129 EISDEG 134

RESULT 54
 HBBC HOPLI STANDARD; PRT; 146 AA.
 AC P82316;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hemoglobin cathodic, beta chain (Hb(Ca) beta chain).
 GN HBB.
 OS Hoplosternum littorale (Hassar).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
 OC Clariichthyidae; Hoplosternum.
 OX NCBI_TaxID=114109;
 RN [1]
 RP SEQUENCE, SUBUNIT, AND MASS SPECTROMETRY.
 RC TISSUE=Blood;
 RX MEDLINE=20298792; PubMed=10747999;
 RA Weber R.E., Pado A., Val A.L., Bang A., Van Hauwaert M.-L.,
 RA Dewilde S., Zaf F., Moens L.;
 RT "Isohemoglobin differentiation in the bimodal-breathing amazon
 catfish Hoplosternum littorale."
 RL J. Biol. Chem. 275:17297-17305(2000).
 CC -1- FUNCTION: Involved in oxygen transport from the lung to the
 various peripheral tissues.
 CC -1- SUBUNIT: Heterotetramer of two alpha chains and two beta chains.
 CC -1- TISSUE SPECIFICITY: Red blood cells.
 CC -1- MASS SPECTROMETRY: MW=15978; MW ERR=0.2; METHOD=Electrospray.
 CC -1- MISCELLANEOUS: This fish has two hemoglobins: cathodic and anodic.
 CC The cathodic Hb and anodic Hb display small and large Bohr effects
 CC respectively. In addition, the cathodic Hb displays a reverse Bohr
 CC effect and appreciable phosphate effects.
 CC -1- SIMILARITY: Belongs to the globin family.

DR HSP; P02142; 1OUT.
 DR InterPro; IPR002337; Beta haem.
 DR InterPro; IPR000971; Globin.
 DR Pfam; PF00042; globin; 1.
 DR PRINTS; PR00814; BETAHAEM.
 DR PROSITE; PS01033; GLOBIN; 1.
 KW Heme; Oxygen transport; Transport; Erythrocyte.
 FT METAL 63 63 IRON (HEME DISTAL LIGAND).
 FT METAL 92 92 IRON (HEME PROXIMAL LIGAND).
 SQ SEQUENCE 146 AA; 15976 MW; 4D75EB9FC8D73539 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 146;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 DAIAAI 40
 DB 9 DAIAAI 14

RESULT 55

Y142 METJA STANDARD; PRT; 147 AA.
 AC Q57607;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein MJ0142.
 GN MJ0142.

OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Kleink H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RL Science 273:1058-1073(1996).
 CC -1- SIMILARITY: STRONG, TO M.JANNASCHII MJ0605.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U67471; AAB98125.1; --
 DR TIGR; MJ0142; --
 DR InterPro; IPR002806; DUF103.
 DR Pfam; PF01953; DUF103; 1.
 DR ProDom; PD015886; DUF103; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 147 AA; 17592 MW; 707D67BF058E065C CRC64;

Query Match 2.8%; Score 6; DB 1; Length 147;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 207 KRELEK 212
 DB 11 KRELEK 16

RESULT 56

DUT THETN STANDARD; PRT; 148 AA.
 AC Q8R346;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Deoxyuridine 5'-triphosphate nucleotidohydrolase (EC 3.6.1.23)
 DE (dUTPase) (dUTP pyrophosphatase).
 GN DUT OR TTE1384.
 OS Thermoanaerobacter tengcongensis.
 OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;

CC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MB4 / JCM 11007;
RA MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RA "A complete sequence of *T. tengcongensis* genome.";
RL Genome Res. 12:689-700(2002).
CC -!- FUNCTION: This enzyme is involved in nucleotide metabolism: it
CC produces dUMP, the immediate precursor of thymidine nucleotides
CC and it decreases the intracellular concentration of dUMP so that
CC uracil cannot be incorporated into DNA (By similarity).
CC -!- CATALYTIC ACTIVITY: dUMP + H(2)O = dUMP + diphosphate.
CC -!- PATHWAY: De novo synthesis of thymidylate.
CC -!- SIMILARITY: Belongs to the dUTPase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE013098; AAM24606.1; -;
DR HAMAP; MF 00116; -; 1.
DR InterPro; IPR008180; DeoxyUTPase.
DR InterPro; IPR008181; dUTPase_1.
DR InterPro; IPR008182; dUTPase_2.
DR Pfam; PF00692; dUTPase; 1.
DR ProDom; PD000946; dUTPase 2; 1.
DR TIGRFAMs; TIGR00576; dut; 1.
KW Hydrolase; Nucleotide metabolism; Complete proteome.
SQ SEQUENCE 148 AA; 16297 MW; A66B21AC647755AA CRC64;

Query Match 2.8%; Score 6; DB 1; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 ANVKGE 95
Db 33 ANVKGE 38
|||||

RESULT 57
YMSW YEAST
ID YMSW YEAST STANDARD; PRT; 150 AA.
AC Q03579;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical 17.2 kDa protein in PRC1-ADE4 intergenic region.
GN YMR298W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c / AB972;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrall B.G.;
RA "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome
RA XIII.";
RL Nature 387:90-93(1997).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X80836; CAAS6807.1; -;
DR PIR; S47459; S47459.
DR Germline; 142975; -;
DR SGD; S0004913; YMR298W.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 21 40 POTENTIAL.
SQ SEQUENCE 150 AA; 17207 MW; 75DBA35225C3065C CRC64;

Query Match 2.8%; Score 6; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LLTAAV 10
Db 32 LLTAAV 37
|||||

RESULT 58
IL3 HYLLA STANDARD; PRT; 152 AA.
ID IL3 HYLLA
AC P06740;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Interleukin-3 precursor (IL-3) (Multipotential colony-stimulating
DE factor) (Hematopoietic growth factor) (P-cell stimulating factor)
DE (Mast-cell growth factor) (MCGF).
GN IL3.
OS Hylobates lar (Common gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=9580;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87002490; PubMed=3489530;
RA Yang Y.-C., Ciavetta A.B., Temple P.A., Chung M.P., Kovacic S.,
RA Witek-Giannotti J.S., Leary A.C., Kriz R., Donahue R.E., Wong G.G.,
RA Clark S.C.;
RA "Human IL-3 (multi-CSF): identification by expression cloning of a
RA novel hematopoietic growth factor related to murine IL-3.";
RL Cell 47:3-10(1986).
CC -!- FUNCTION: GRANULOCYTE/MACROPHAGE COLONY-STIMULATING FACTORS ARE
CC CYTOKINES THAT ACT IN HEMATOPOIESIS BY CONTROLLING THE PRODUCTION,
CC DIFFERENTIATION, AND FUNCTION OF 2 RELATED WHITE CELL POPULATIONS
CC OF THE BLOOD, THE GRANULOCYTES AND THE MONOCYTES-MACROPHAGES.
CC -!- FUNCTION: THIS CSF INDUCES GRANULOCYTES, MACROPHAGES, MAST CELLS,
CC STEM CELLS, ERYTHROID CELLS, EOSINOPHILS AND MEGAKARYOCYTES.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Activated T cells, mast cells, natural killer
CC cells.
CC -!- SIMILARITY: Belongs to the IL-3 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14744; AAA35455.1; -;
DR HSSP; P08700; IL3I.
DR InterPro; IPR002183; Interleukin_3.
DR Pfam; PF02059; IL3; 1.
DR PRINTS; PR00430; INTERLEUKIN3.


```

DR InterPro; IPR003815; LuxS.
DR Pfam; PF02664; LuxS; 1.
DR PRINTS; PR01487; LUXSPROTEIN.
DR ProDom; PD013172; LuxS; 1.
KW Quorum sensing; Autoinducer synthesis; Iron; Metal-binding;
KW Hydrolase; Complete proteome.
FT METAL 54 54 IRON (BY SIMILARITY).
FT METAL 58 58 IRON (BY SIMILARITY).
FT METAL 125 125 IRON (BY SIMILARITY).
SQ SEQUENCE 158 AA; 17605 MW; 838239F57545664A CRC64;

Query Match 2.8%; Score 6; DB 1; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 TKVLAP 204
Db 12 TKVLAP 17

RESULT 61
LUXS_LACPL STANDARD; PRT; 158 AA.
ID LUXS_LACPL STANDARD; PRT; 158 AA.
AC Q88YI6;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE S-ribosylhomocysteinease (EC 3.13.1.-) (Autoinducer-2 production
DE protein LuxS) (AI-2 synthesis protein).
DE LUXS OR LP 0774.
GN LUXS
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1590;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=NCIMB 8826 / WCFS1;
RX MEDLINE=22480296; PubMed=12566566;
RA Kleerebezem M., Bokhorst J., van Kranenburg R., Molenaar D.,
RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
RA Piers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
RA De Vos W.M., Siezen R.J.;
RT "Complete genome sequence of Lactobacillus plantarum WCFS1."
CC Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
CC -!- FUNCTION: Involved in the synthesis of autoinducer 2 (AI-2) which
CC is secreted by bacteria and is used to communicate both the cell
CC density and the metabolic potential of the environment. The
CC regulation of gene expression in response to changes in cell
CC density is called quorum sensing. Catalyzes the transformation of
CC S-ribosylhomocysteine (RHC) to homocysteine (HC) and 4,5-
CC dihydroxy-2,3-pentadione (DPPD) (By similarity).
CC -!- CATALYTIC ACTIVITY: S-ribosylhomocysteine + H(2)O = homocysteine +
CC 4,5-dihydroxy-2,3-pentanedione.
CC -!- COFACTOR: Iron (By similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SIMILARITY: Belongs to the luxS family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U24099; AAC50151.1; -.
CC PIR; I61900; I61900.
CC HSSP; P00656; ILSQ.
CC InterPro; IPR001427; RNaseA.
CC Pfam; PF00074; RNaseA; 1.
CC PRINTS; PR00794; RIBONUCLEASE.
CC ProDom; PD000535; RNaseA; 1.
CC SMART; SM00092; RNase P; 1.
CC PROSITE; PS00127; RNASE_PANCREATIC; 1.
KW Hydrolase; Nuclease; Endonuclease; Eosinophil; Glycoprotein; Signal.
FT SIGNAL 1 27 BY SIMILARITY.
FT CHAIN 28 158 NONSECRETORY RIBONUCLEASE.
FT ACT_SITE 42 42 BY SIMILARITY.
FT ACT_SITE 65 65 BY SIMILARITY.
FT ACT_SITE 153 153 BY SIMILARITY.
FT DISULFID 50 110 BY SIMILARITY.
FT DISULFID 64 121 BY SIMILARITY.
FT DISULFID 82 136 BY SIMILARITY.
FT DISULFID 89 98 BY SIMILARITY.
FT CARBOHYD 34 34 C-LINKED (MAN) (BY SIMILARITY).

```



```
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 92 92 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 158 AA; 1770 MW; 88864C77806EBE5C CRC64;

Query Match 2.8%; Score 6; DB 1; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 FANVVN 167
Db 75 FANVVN 80

RESULT 63
SSRP THETN STANDARD; PRT; 158 AA.
AC QRB39;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SsrA-binding protein.
DE SsrA-binding protein.
GN SmpB OR TTE0985
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RA "A complete sequence of T. tengcongensis genome.";
RA Genome Res. 12:689-700(2002).
CC -!- FUNCTION: Binds specifically to the ssrA RNA (tmRNA) and is
CC required for stable association of ssrA with ribosomes (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Belongs to the smpB family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AE013063; AAM24240.1; -
DR HAMAP: MF_00023; -; 1.
DR InterPro: IPR000037; SmpB.
DR Pfam: PF01669; SmpB; 1.
DR ProDom: PD004488; SmpB; 1.
DR TIGRFAMs: TIGR00086; smpB; 1.
DR PROSITE: PS01317; SSRP; 1.
DR RNA-binding: Complete proteome.
SQ SEQUENCE 158 AA; 18530 MW; E7C66787465941D4 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 207 KRELEK 212
Db 144 KRELEK 149

RESULT 64
SSRP ECOLI STANDARD; PRT; 159 AA.
AC P32052; P77011;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SsrA-binding protein.
DE SsrA-binding protein.
GN SmpB OR TTE0985
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RA "A complete sequence of T. tengcongensis genome.";
RA Genome Res. 12:689-700(2002).
CC -!- FUNCTION: Binds specifically to the ssrA RNA (tmRNA) and is
CC required for stable association of ssrA with ribosomes (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Belongs to the smpB family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AE013063; AAM24240.1; -
DR HAMAP: MF_00023; -; 1.
DR InterPro: IPR000037; SmpB.
DR Pfam: PF01669; SmpB; 1.
DR ProDom: PD004488; SmpB; 1.
DR TIGRFAMs: TIGR00086; smpB; 1.
DR PROSITE: PS01317; SSRP; 1.
DR RNA-binding: Complete proteome.
SQ SEQUENCE 158 AA; 18530 MW; E7C66787465941D4 CRC64;
```

```
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SsrA-binding protein (Small protein B).
GN SmpB OR SMOB OR B2620 OR C3142 OR C3913 OR ECG3482 OR SF2679 OR S2857.
OS Escherichia coli.
OS Escherichia coli O6.
OS Escherichia coli O157:H7, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 217992, 83334, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=95023883; PubMed=7524073;
RA Komine Y., Kitabatake M., Yokogawa T., Nishikawa K., Inokuchi H.;
RA "A tRNA-like structure is present in 108A RNA, a small stable RNA
RA from Escherichia coli.";
RA Proc. Natl. Acad. Sci. U.S.A. 91:9223-9227(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12.";
RA Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
RA Oshima T., Oyama S., Saito N., Sampaio G., Sato Y., Sivasubram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RA "Construction of a contiguous 874-kb sequence of the Escherichia coli
RA K-12 genome corresponding to 50.0-68.8 min on the linkage map and
RA analysis of its sequence features.";
RA DNA Res. 4:91-113(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA McElroy H.L.T., Donnenberg M.S., Blattner F.R.;
RA "Extensive mosaic structure revealed by the complete genome sequence
RA of uropathogenic Escherichia coli.";
RA Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RA Nature 409:529-533(2001).
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
```


"Eosinophil cationic protein cDNA. Comparison with other toxic cationic proteins and ribonucleases.";
J. Immunol. 143:952-955(1989).
[4]
SEQUENCE FROM N.A.
TISSUE=Colon;
RA Simonsen C.C., Kennedy J., Comstock L., Ashton N., McGrogan M.;
RL Submitted (OCT-1990) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE FROM N.A., AND VARIANT THR-124.
RX MEDLINE=20556158; PubMed=11102386;
RA Zhang J., Rosenberg H.F.;
RL "Sequence variation at two eosinophil-associated ribonuclease loci in humans.";
RL Genetics 156:1949-1958(2000).
[6]
SEQUENCE OF 20-160 FROM N.A., AND VARIANTS CYS-72 AND THR-124.
RA Eystrom J., Molin D., Jonsson U.B., Enblad G., Sundstrom C.,
RA Hogbom E., Verge P.;
RL "Identification of polymorphisms in the ECP gene. Relation to disease activity in Hodgkins lymphoma.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
[7]
SEQUENCE OF 28-87.
RX MEDLINE=86205853; PubMed=3458170;
RA Gleich G.J., Loegering D.A., Bell M.P., Checkel J.L., Ackerman S.J.,
RA McKean D.J.;
RL "Biochemical and functional similarities between human eosinophil-derived neurotoxin and eosinophil cationic protein: homology with ribonuclease.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:3146-3150(1986).
[8]
SEQUENCE OF 28-47, AND ANTIMICROBIAL ACTIVITY.
RX MEDLINE=89315847; PubMed=2501794;
RA Gabay J.E., Scott R.W., Campanelli D., Griffith J., Wilde C.,
RA Marra M.N., Seeger M., Nathan C.F.;
RL "Antibiotic proteins of human polymorphonuclear leukocytes.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:5610-5614(1989).
[9]
X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
RX MEDLINE=20366247; PubMed=10903870;
RA Mallorqui-Fernandez G., Pous J., Peracaula R., Aymami J., Maeda T.,
RA Tada H., Yamada H., Seno M., de Llorens R., Gomis-Ruth F.X., Coll M.;
RL "Three-dimensional crystal structure of human eosinophil cationic protein (RNase 3) at 1.75 A resolution.";
RL J. Mol. Biol. 300:1297-1307(2000).
[10]
X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=22243790; PubMed=12356310;
RA Mohan C.G., Boix E., Evans H.R., Nikolovski Z., Noguez M.V.,
RA Cuchillo C.M., Acharya K.R.;
RL "The crystal structure of eosinophil cationic protein in complex with 2',5'-ADP at 2.0 A resolution reveals the details of the ribonucleolytic active site.";
RL Biochemistry 41:12100-12106(2002).
CC -!- FUNCTION: Cytotoxin and helminthotoxin with low-efficiency ribonuclease activity. Possesses a wide variety of biological activities. Exhibits antibacterial activity.
CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific granule.
CC -!- DISEASE: Induces the neurotoxic effect known as the Gordon phenomenon.
CC -!- SIMILARITY: Belongs to the pancreatic ribonuclease family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X15161; CAA33251.1; -

EMBL; X16545; CAA34545.1; -
EMBL; M28128; AAA50283.1; -
EMBL; X55990; CAA39462.1; -
EMBL; AF294019; AAG31589.1; -
EMBL; AF294020; AAG31590.1; -
EMBL; AF294021; AAG31591.1; -
EMBL; AF294022; AAG31592.1; -
EMBL; AF294023; AAG31593.1; -
EMBL; AF294024; AAG31594.1; -
EMBL; AF294025; AAG31595.1; -
EMBL; AF294026; AAG31596.1; -
EMBL; AF441204; AAL35279.1; -
EMBL; AF441205; AAL35280.1; -
EMBL; AF441206; AAL35281.1; -
EMBL; B35328; J10106.
PDB; 1DHT; 08-FEB-01.
PDB; 1H1H; 03-OCT-02.
PDB; 1QMT; 04-FEB-00.
Genew; KGC:10046; RNASE3.
MIM; 131398; -
GO; GO:000576; C:extracellular; TAS.
GO; GO:0005803; C:secretory vesicle; TAS.
GO; GO:0005625; C:soluble fraction; TAS.
GO; GO:0004540; F:ribonuclease activity; TAS.
GO; GO:0006401; P:RNA catabolism; TAS.
InterPro; IPR001427; RNaseA.
Pfam; PF00074; rnasea; 1.
PRINTS; PR00794; RIBONUCLEASE.
PRODOM; PD000535; RNaseA; 1.
SMART; SM00092; RNase PC; 1.
PROSITE; PS00127; RNASE PANCREATIC; 1.
Hydrolase; Nuclease; Endonuclease; Eosinophil; Glycoprotein;
Antibiotic; Signal; Polymorphism; 3D-structure.
FT SIGNAL 1 27 EOSINOPHIL CATIONIC PROTEIN.
FT CHAIN 28 160
FT ACT_SITE 42 42
FT ACT_SITE 65 65
FT ACT_SITE 155 155
FT DISULFID 50 110
FT DISULFID 64 123
FT DISULFID 82 138
FT DISULFID 89 98
FT CARBOHYD 84 84
FT CARBOHYD 92 92
FT CARBOHYD 119 119
FT VARIANT 72 72 /FTID=VAR 014109.
FT VARIANT 124 124 R -> T (in dbSNP:2073342).
FT SEQUENCE 160 AA; 18440 MW; D7BED24F62486009 CRC64;
SQ
Query Match 2.8%; Score 6; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 162 FANVW 167
Db 75 FANVW 80
RESULT 68
ECP_PANTR
ID_ECP_PANTR STANDARD; PRT; 160 AA.
AC P47780;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Eosinophil cationic protein precursor (EC 3.1.27.-) (ECP)
DE (Ribonuclease 3) (RNase 3).
GN RNASE3 OR RN3.
OS Pan troglodytes (Chimpanzee).
OC Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

```
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9532576; PubMed=7663519;
RA Rosenberg H.F., Dyer K.D., Tiffany H.L., Gonzalez M.;
RT "Rapid evolution of a unique family of primate ribonuclease genes.";
RL Nat. Genet. 10:219-223(1995).
CC -!- FUNCTION: Cytotoxin and helminthotoxin with low-efficiency
CC ribonuclease activity. Possesses a wide variety of biological
CC activities. Exhibits antibacterial activity (By similarity).
CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific
CC granule.
CC -!- SIMILARITY: Belongs to the pancreatic ribonuclease family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U24103; AAC50150.1; -.
CC PIR; I61899; I61899.
CC HSP; P00656; 7RSA.
CC InterPro; IPR001427; RNaseA.
CC Pfam; PF00074; rnaaseA; 1.
CC PRINTS; PR00794; RIBONUCLEASE.
CC ProDom; PD000535; RNaseA; 1.
CC SMART; SM00092; RNaseA; 1.
CC PROSITE; PS00127; RNASE_PANCREATIC; 1.
CC Hydroxylase; Nuclease; Eosinophil; Glycoprotein; Signal.
CC SIGNAL 1 27
CC CHAIN 28 160
CC ACT_SITE 42 42
CC ACT_SITE 65 65
CC ACT_SITE 155 155
CC ACT_SITE 50 110
CC DISULFID 64 123
CC DISULFID 82 138
CC DISULFID 89 98
CC CARBOHYD 84 84
CC CARBOHYD 92 92
CC CARBOHYD 119 119
CC SEQUENCE 160 AA; 18341 MW; C6DFFAEFD650A41 CRC64;
Query Match 2.8%; Score 6; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 162 FANVNV 167
Db 75 FANVNV 80
RESULT 69
ECP_PONFY STANDARD; PRT; 160 AA.
AC P47781;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Eosinophil cationic protein precursor (EC 3.1.27.-) (ECP)
DE (Ribonuclease 3) (RNase 3).
GN RNase3 OR RNS3.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9532576; PubMed=7663519;
RA Rosenberg H.F., Dyer K.D., Tiffany H.L., Gonzalez M.;
RT "Rapid evolution of a unique family of primate ribonuclease genes.";
RL Nat. Genet. 10:219-223(1995).
CC -!- FUNCTION: Cytotoxin and helminthotoxin with low-efficiency
CC ribonuclease activity. Possesses a wide variety of biological
CC activities. Exhibits antibacterial activity (By similarity).
CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific
CC granule.
CC -!- SIMILARITY: Belongs to the pancreatic ribonuclease family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U24103; AAC50150.1; -.
CC PIR; I61899; I61899.
CC HSP; P00656; 7RSA.
CC InterPro; IPR001427; RNaseA.
CC Pfam; PF00074; rnaaseA; 1.
CC PRINTS; PR00794; RIBONUCLEASE.
CC ProDom; PD000535; RNaseA; 1.
CC SMART; SM00092; RNaseA; 1.
CC PROSITE; PS00127; RNASE_PANCREATIC; 1.
CC Hydroxylase; Nuclease; Eosinophil; Glycoprotein; Signal.
CC SIGNAL 1 27
CC CHAIN 28 160
CC ACT_SITE 42 42
CC ACT_SITE 65 65
CC ACT_SITE 155 155
CC ACT_SITE 50 110
CC DISULFID 64 123
CC DISULFID 82 138
CC DISULFID 89 98
CC CARBOHYD 84 84
CC CARBOHYD 92 92
CC CARBOHYD 119 119
CC SEQUENCE 160 AA; 18341 MW; C6DFFAEFD650A41 CRC64;
```

```
RT "Rapid evolution of a unique family of primate ribonuclease genes.";
RL Nat. Genet. 10:219-223(1995).
CC -!- FUNCTION: Cytotoxin and helminthotoxin with low-efficiency
CC ribonuclease activity. Possesses a wide variety of biological
CC activities. Exhibits antibacterial activity (By similarity).
CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific
CC granule.
CC -!- SIMILARITY: Belongs to the pancreatic ribonuclease family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U24101; AAC50147.1; -.
CC PIR; I61896; I61896.
CC HSP; P00656; 1LSQ.
CC InterPro; IPR001427; RNaseA.
CC Pfam; PF00074; rnaaseA; 1.
CC PRINTS; PR00794; RIBONUCLEASE.
CC ProDom; PD000535; RNaseA; 1.
CC SMART; SM00092; RNaseA; 1.
CC PROSITE; PS00127; RNASE_PANCREATIC; 1.
CC Hydroxylase; Nuclease; Eosinophil; Glycoprotein; Signal.
CC SIGNAL 1 27
CC CHAIN 28 160
CC ACT_SITE 42 42
CC ACT_SITE 65 65
CC ACT_SITE 155 155
CC ACT_SITE 50 110
CC DISULFID 64 123
CC DISULFID 82 138
CC DISULFID 89 98
CC CARBOHYD 86 86
CC CARBOHYD 92 92
CC CARBOHYD 111 111
CC CARBOHYD 119 119
CC SEQUENCE 160 AA; 18262 MW; 1B1A70A85C0A2BB4 CRC64;
Query Match 2.8%; Score 6; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 162 FANVNV 167
Db 75 FANVNV 80
RESULT 70
SSRP_HAEDU STANDARD; PRT; 160 AA.
AC Q7VM64;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE SsrA-binding protein.
DE SMPB OR HD1134.
OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=730;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=35000HP / ATCC 700724;
RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
RT "The complete genome sequence of Haemophilus ducreyi."
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Binds specifically to the ssrA RNA (tmRNA) and is
CC required for stable association of ssrA with ribosomes (By
```

```
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Belongs to the smpB family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE017153; AAP95996.1; -.
DR HAMAP; MF_00023; -.
DR InterPro; IPR000037; SmpB.
DR Pfam; PF01668; SmpB; 1.
DR ProDom; PDO04488; SmpB; 1.
DR TIGRFAMs; TIGR00086; SmpB; 1.
DR PROSITE; PS01317; SSRP; 1.
KW RNA-binding; Complete proteome.
SQ SEQUENCE 160 AA; 18098 MW; B082907668CBDDCD CRC64;

Query Match 2.8%; Score 6; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVALSL 144
DB 109 VVALSL 114

RESULT 71
SSRP PHOHL
ID SSRP PHOHL STANDARD; PRT; 160 AA.
AC Q7NU1.
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE SsrA-binding protein.
GN SMPB OR PU3378.
OS Photobacterium luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photobacterius.
OX NCBI_TaxID=141679;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TT01.
RX MEDLINE=22957627; PubMed=14528314;
RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taurit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,
RA Dassa E., Derosé R., Derzelle S., Freysinet G., Gaudriault S.,
RA Médigue C., Lanols A., Powell K., Sigulier P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photobacterius
RT luminescens".
RL Nat. Biotechnol. 21:1307-1313(2003).
CC -!- FUNCTION: Binds specifically to the ssrA RNA (tmRNA) and is
CC required for stable association of ssrA with ribosomes (By
CC similarity).
CC -!- SIMILARITY: Belongs to the smpB family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BX571870; CAE15752.1; -.
DR HAMAP; MF_00023; -.
DR
```

```
DR PROSITE; PS01317; SSRP; 1.
KW RNA-binding; Complete proteome.
SQ SEQUENCE 160 AA; 18224 MW; 4BAEB9C8BF0B15D CRC64;

Query Match 2.8%; Score 6; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVALSL 144
DB 111 VVALSL 116

RESULT 72
YQGC_BACSU
ID YQGC_BACSU STANDARD; PRT; 160 AA.
AC P54486;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein yqgC.
GN YQGC OR BSU25030.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=97124195; PubMed=8969508;
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
RA Kobayashi Y.;
RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
RT the Bacillus subtilis genome containing the skin element and many
RT sporulation genes".
RL Microbiology 142:3103-3111(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessières P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Chim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kashara Y., Koningstein G., Krogh S., Kumano M.,
RA Kobayashi Y., Kottter P., Koningstein G., Lauber J., Lazarevic V.,
RA Kurita K., Lapidus A., Lardinot S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue J.C., Médigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Pressac E., Fujic P., Fumelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Sertor P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni K.,
RA Tosaro V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D84432; BAA12506.1; --
DR EMBL; Z99116; CAB14433.1; --
DR PIR; G69955; G69955.
DR Subtilisin; BG11675; Y9GC.
DR InterPro; IPR007403; DUF456.
DR Pfam; PF04306; DUF456; 1.
DR Hypothetical protein; Transmembrane; Complete proteome.
KW TRANSNEM 5 25 POTENTIAL.
FT TRANSNEM 28 48 POTENTIAL.
FT TRANSNEM 49 69 POTENTIAL.
FT TRANSNEM 89 109 POTENTIAL.
FT TRANSNEM 137 157 POTENTIAL.
SQ SEQUENCE 160 AA; 17295 MW; EE7B9346D75B37EE CRC64;

Query Match 2.8%; Score 6; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LLIAAV 10
Db 7 LLIAAV 12
|||||

RESULT 73
RKND GORGO STANDARD; PRT; 161 AA.
ID RKND GORGO STANDARD; PRT; 161 AA.
AC P47782;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nonsecretory ribonuclease precursor (EC 3.1.27.5) (Ribonuclease US)
DE (Eosinophil-derived neurotoxin) (RNase Upi-2) (Ribonuclease 2)
DE (RNase 2).
GN RNASE2 OR RNS2.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9595;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95392576; PubMed=7663519;
RA Rosenberg H.F., Dyer K.D., Tiffany H.L., Gonzalez M.;
RT "Rapid evolution of a unique family of primate ribonuclease genes.";
RL Nat. Genet. 10:219-223(1995).
CC -1- FUNCTION: This is a non-secretory ribonuclease. It is a pyrimidine
CC specific nuclease with a slight preference for U. Cytotoxin and
CC helminthotoxin. Possesses a wide variety of biological activities.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to nucleoside 3'-
CC phosphates and 3'-phosphooligonucleotides ending in C-P or U-P
CC with 2',3'-cyclic phosphate intermediates.
CC -1- SUBCELLULAR LOCATION: Lysosomal (Probable); matrix of eosinophil's
CC large specific granule.
CC -1- SIMILARITY: Belongs to the pancreatic ribonuclease family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U24100; AAC50144.1; --
DR PIR; I37034; I37034.
DR HSSP; P00656; ILISQ.
DR InterPro; IPR001427; RNaseA.

DR Pfam; PF00074; rnaseA; 1.
DR PRINTS; PR00794; RIBONUCLEASE.
DR ProDom; PD000535; RNaseA; 1.
DR SMART; SM00092; RNase P; 1.
DR PROSITE; PS00127; RNASE_PANCREATIC; 1.
KW Hydrolase; Nuclease; Endonuclease; Eosinophil; Glycoprotein; Signal.
FT SIGNAL 1 27 BY SIMILARITY.
FT CHAIN 28 161 NONSECRETORY RIBONUCLEASE.
FT ACT SITE 42 42 BY SIMILARITY.
FT ACT SITE 65 65 BY SIMILARITY.
FT ACT SITE 156 156 BY SIMILARITY.
FT DISULFID 50 110 BY SIMILARITY.
FT DISULFID 64 123 BY SIMILARITY.
FT DISULFID 82 138 BY SIMILARITY.
FT DISULFID 89 98 BY SIMILARITY.
FT CARBOHYD 34 34 C-LINKED (MAN) (BY SIMILARITY).
FT CARBOHYD 44 44 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 92 92 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 111 111 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 161 AA; 18382 MW; 0D2257A950555F04 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 FANVNV 167
Db 75 FANVNV 80
|||||

RESULT 74
RKND HUMAN STANDARD; PRT; 161 AA.
ID RKND HUMAN STANDARD; PRT; 161 AA.
AC P10153; QOH2B7;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Nonsecretory ribonuclease precursor (EC 3.1.27.5) (Ribonuclease US)
DE (Eosinophil-derived neurotoxin) (RNase Upi-2) (Ribonuclease 2)
DE (RNase 2).
GN RNASE2 OR RNS2 OR EDN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89282792; PubMed=2734298;
RA Rosenberg H.F., Tenen D.G., Ackerman S.J.;
RT "Molecular cloning of the human eosinophil-derived neurotoxin: a
RT member of the ribonuclease gene family.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4460-4464(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90076963; PubMed=2591744;
RA Hamann K.J., Barker R.L., Loegering D.A., Pease L.R., Gleich G.J.;
RT "Sequence of human eosinophil-derived neurotoxin cDNA: identity of
RT deduced amino acid sequence with human nonsecretory ribonucleases.";
RL Gene 83:161-167(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=90353951; PubMed=2387583;
RA Hamann K.J., Ten R.M., Loegering D.A., Jenkins R.B., Heise M.T.,
RA Schad C.R., Pease L.R., Gleich G.J., Barker R.L.;
RT "Structure and chromosome localization of the human eosinophil-derived
RT neurotoxin and eosinophil cationic protein genes: evidence for
RT intronless coding sequences in the ribonuclease gene superfamily.";
RL Genomics 7:535-546(1990).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon, and Leukemia;

RA Simonsen C.C., Kennedy J., Comstock L., Ashton N., McGrogan M.;
 RL Submitted (OCT-1990) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89309773; PubMed=2745977;
 RA Barker R.L., Loegering D.A., Ten R.M., Hamann K.J., Pease L.R.,
 RA Gleich G.J.;
 RT "Eosinophil cationic protein cDNA. Comparison with other toxic
 RT cationic proteins and ribonucleases.";
 RL J. Immunol. 143:952-955 (1989).
 RN [6]
 RP SEQUENCE FROM N.A., AND VARIANT ASN-156.
 RX MEDLINE=20556158; PubMed=1102386;
 RA Zhang J., Rosenberg H.F.;
 RT "Sequence variation at two eosinophil-associated ribonuclease loci in
 RT humans.";
 RL Genetics 156:1949-1958 (2000).
 RN [7]
 RP SEQUENCE OF 28-161.
 RX MEDLINE=89006644; PubMed=3166997;
 RA Beintema J.J., Hofsteenge J., Iwama M., Morita T., Ohgi K., Irie M.,
 RA Sugiyama R.H., Schieven G.L., Dekker C.A., Glitz D.G.;
 RT "Amino acid sequence of the nonsecretory ribonuclease of human
 RT urine.";
 RL Biochemistry 27:4530-4538 (1988).
 RN [8]
 RP SEQUENCE OF 28-55.
 RX MEDLINE=85261183; PubMed=326759;
 RA Niwata Y., Ohgi K., Sada A., Takizawa Y., Irie M.;
 RT "Purification and properties of bovine kidney ribonucleases.";
 RL J. Biochem. 97:923-934 (1985).
 RN [10]
 RP SEQUENCE OF 28-53.
 RC TISSUE=Liver;
 RX MEDLINE=89034073; PubMed=3182786;
 RA Sorrentino S., Tucker G.K., Glitz D.G.;
 RT "Purification and characterization of a ribonuclease from human
 RT liver.";
 RL J. Biol. Chem. 263:16125-16131 (1988).
 RN [11]
 RP SEQUENCE OF 24-43.
 RC TISSUE=Urine;
 RX MEDLINE=92268031; PubMed=1587793;
 RA Sakakibara R., Hashida K., Kitahara T., Ishiguro M.;
 RT "Characterization of a unique nonsecretory ribonuclease from urine of
 RT pregnant women.";
 RL J. Biochem. 111:325-330 (1992).
 RN [12]
 RP INVOLVEMENT IN CHEMOTAXIS.
 RX MEDLINE=22929839; PubMed=12855582;
 RA Yang D., Rosenberg H.F., Chen Q., Dyer K.D., Kurosaka K.,
 RA Oppenheim J.J.;
 RT "Eosinophil-derived neurotoxin (EDN), an antimicrobial protein with
 RT chemotactic activities for dendritic cells.";
 RL Blood 102:3396-3403 (2003).
 RN [13]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE=96187458; PubMed=8710835;
 RA Church W.B., Palmer A., Wathey J.C., Kitson D.H.;
 RT "Homology modeling of histidine-containing phosphocarrier protein and
 RT eosinophil-derived neurotoxin: construction of models and comparison
 RT with experiment.";
 RL Proteins 23:422-430 (1995).
 RN [14]
 RP X-RAY CRYSTALLOGRAPHY (1.83 ANGSTROMS).
 RX MEDLINE=96319750; PubMed=8759319;
 RA Mosmann S.C., Newton D.L., Youle R.J., James M.N.G.;
 RT "X-ray crystallographic structure of recombinant eosinophil-derived
 RT neurotoxin at 1.83-A resolution.";
 RL J. Mol. Biol. 260:540-552 (1996).
 RN [15]
 RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
 RX MEDLINE=21226742; PubMed=11154698;
 RA Leonidas D.D., Boix E., Prill R., Suzuki M., Turton R., Minson K.,
 RA Swaminathan G.J., Youle R.J., Acharya K.R.;
 RT "Mapping the ribonucleolytic active site of eosinophil-derived
 RT neurotoxin (EDN). High resolution crystal structures of EDN complexes
 RT with adenylic nucleotide inhibitors.";
 RL J. Biol. Chem. 276:15009-15017 (2001).
 RN [16]
 RP CARBOHYDRATE-LINKAGE SITE TRP-34.
 RX MEDLINE=95034787; PubMed=7947762;
 RA Hofsteenge J., Mueller D.R., de Beer T., Loeffler A., Richter W.J.,
 RA Vilgertshart J.F.G.;
 RT "New type of linkage between a carbohydrate and a protein: C-
 RT glycosylation of a specific tryptophan residue in human RNase Us.";
 RL Biochemistry 33:13524-13530 (1994).
 RN [17]
 RP STRUCTURE OF C-GLYCOSYLATED GROUP.
 RX MEDLINE=96018866; PubMed=7547911;
 RA de Beer T., Vilgertshart J.F.G., Loeffler A., Hofsteenge J.;
 RT "The hexopyranosyl residue that is C-glycosidically linked to the side
 RT chain of tryptophan-7 in human RNase Us is alpha-mannopyranose.";
 RL Biochemistry 34:11785-11789 (1995).
 RN [18]
 RP RECOGNITION SITE FOR C-GLYCOSYLATION.
 RX MEDLINE=98119743; PubMed=9450956;
 RA Krieg J., Hartmann S., Vicentini A., Glasner W., Hess D.,
 RA Hofsteenge J.;
 RT "Recognition signal for C-mannosylation of Trp-7 in RNase 2 consists
 RT of sequence Trp-x-Trp.";
 RL Mol. Biol. Cell 9:301-309 (1998).
 CC -!- FUNCTION: This is a non-secretory ribonuclease. It is a pyrimidine
 CC specific nuclease with a slight preference for U. Cytotoxin and
 CC helminthotoxin. Selectively chemotactic for dendritic cells.
 CC -!- Possesses a wide variety of biological activities.
 CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to nucleoside 3'-
 CC phosphates and 3'-phosphogluconolactones ending in C-P or U-P
 CC with 2',3'-cyclic phosphate intermediates.
 CC -!- SUBCELLULAR LOCATION: Lysosomal (Probable); matrix of eosinophil's
 CC large specific granule.
 CC -!- TISSUE SPECIFICITY: Liver, lung, spleen, leukocytes and body
 CC fluids.
 CC -!- DOMAIN: The N-terminal region is necessary for mediating
 CC chemotactic activity.
 CC -!- PTM: A particular signal processing and glycosylation pattern may
 CC differentiate the Upi2 RNase, found specifically in pregnant women
 CC urine, from other nonsecretory RNases.
 CC -!- DISEASE: EDN induces ataxia and paralysis, the neurotoxic effect
 CC known as the Gordon phenomenon.
 CC -!- SIMILARITY: Belongs to the pancreatic ribonuclease family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M28129; AAA50284.1; -
 CC EMBL: X16546; CAA34546.1; -
 CC EMBL: M24157; AAA52337.1; -
 CC EMBL: M30510; AAC82505.1; -
 CC EMBL: X55987; CAA39459.1; -
 CC EMBL: X55988; CAA39460.1; -
 CC EMBL: AF294007; AAG31577.1; -

DR	EMBL; AF294016; AAG31586.1; -.
DR	EMBL; AF294017; AAG31587.1; -.
DR	PIR; I61898; ILSO.
DR	HSP; P00656; RNaseA.
DR	InterPro; IPRO01427; RNaseA.
DR	Pfam; PF00074; RNaseA_1
DR	PRINTS; PR00794; RIBONUCLEASE.
DR	ProDom; PD000535; RNaseA; 1.
DR	SMART; SN00092; RNase_Pc; 1.
DR	PROSITE; PS00127; RNASE PANCREATIC; 1.
DR	KW Hydrolase; Nuclease; Endonuclease; Eosinophil; Glycoprotein; Signal.
FT	SIGNAL 1 27 BY SIMILARITY.
FT	CHAIN 28 161 NONSECRETORY RIBONUCLEASE.
FT	ACT_SITE 42 42 BY SIMILARITY.
FT	ACT_SITE 65 65 BY SIMILARITY.
FT	ACT_SITE 156 156 BY SIMILARITY.
FT	CARBOHYD 34 34 C-LINKED (MAN) (BY SIMILARITY).
FT	DISULFID 50 110 BY SIMILARITY.
FT	DISULFID 64 123 BY SIMILARITY.
FT	DISULFID 82 138 BY SIMILARITY.
FT	DISULFID 89 98 BY SIMILARITY.
FT	CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 92 92 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQ	SEQUENCE 161 AA; 19345 MW; 8321P4596CBFP938 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 161;
 Best Local Similarity 100.0%; Pred.No.1.5e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	162 FANVNVN 167
DB	 75 FANVNVN 80

RESULT 76
 RNKD PONPY ID RNKD PONPY STANDARD; PRT; 161 AA.
 AC P47784;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Nonsecretory ribonuclease precursor (EC 3.1.27.5) (Ribonuclease US)
 DE (Eosinophil-derived neurotoxin) (RNase Upi-2) (Ribonuclease 2)
 DE (Enase 2).
 GN RNASB2 OR RNS2.
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
 NCBI_TaxId=9600;
 RP [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=9532576; PubMed=7663519;
 RA Rosenberg H.F., Dyer K.D., Tiffany H.L., Gonzalez M.;
 RT "Rapid evolution of a unique family of primate ribonuclease genes.";
 RL Nat. Genet. 10:219-223(1995).
 CC -I- FUNCTION: This is a non-secretory ribonuclease. It is a pyrimidine specific nuclease with a slight preference for U. Cytotoxin and helminthotoxin. Possesses a wide variety of biological activities.
 CC -I- CATALYTIC ACTIVITY: Endonucleolytic cleavage to nucleoside 3'-phosphates and 3'-phosphooligonucleotides ending in C-P or U-P with 2',3'-cyclic phosphate intermediates.
 CC -I- SUBCELLULAR LOCATION: Lysosomal (Probable); matrix of eosinophil's large specific granule.
 CC -I- SIMILARITY: Belongs to the pancreatic ribonuclease family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial purposes without permission from the copyright owner is prohibited.

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U24104; AAC50148.1; -.
CC PIR; I61897; I61897.
CC HSP; P00656; IAZW.
CC InterPro: IPR001427; RNaseA.
CC Pfam; PF00074; RNaseA.
CC PRINTS; PR00794; RIBONUCLEASE.
CC ProDom; PD000535; RNaseA; 1.
CC SMART; SM00092; RNaseA; 1.
CC PROSITE; PS00127; RNASE PANCREATIC; 1.
CC Hydrolase; Nuclease; Endonuclease; Boinophil; Glycoprotein; Signal.
FT SIGNAL 1 27
FT CHAIN 28 161
FT ACT_SITE 42 42
FT ACT_SITE 65 65
FT ACT_SITE 156 156
FT DISULFID 50 110
FT DISULFID 64 123
FT DISULFID 82 138
FT DISULFID 89 98
FT CARBOHYD 34 34
FT CARBOHYD 44 44
FT CARBOHYD 86 86
FT CARBOHYD 92 92
FT CARBOHYD 111 111
FT CARBOHYD 119 119
FT SEQUENCE 161 AA; 18327 MW; P237D6D1566750C5 CRC64;
Query Match 2.8%; Score 6; DB 1; Length 161;
Best Local Similarity 100.0%; Pred.No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 FANVNV 167
DB 75 FANVNV 80

RESULT 77
NOS2 MACMU
ID NOS2 MACMU STANDARD; PRT; 162 AA.
AC O46650;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Nitric oxide synthase, inducible (EC 1.14.13.39) (NOS, type II)
DE (Inducible NOS) (iNOS) (Fragment).
OS NOS2.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Microglia;
RX MEDLINE=97056192; PubMed=8900532;
RA Lane T.E., Buchmeier M.J., Wary D.D., Fox H.S.;
RT "Expression of inflammatory cytokines and inducible nitric oxide
RT synthase in brains of HIV-infected thymus monkeys: applications to
RT HIV-induced central nervous system disease.";
RL Mol. Med. 2:27-37(1996).
CC -!- FUNCTION: Produces nitric oxide (NO) which is a messenger molecule
CC with diverse functions throughout the body. iNOS and NO may
CC contribute to sinian immunodeficiency virus (SIV)-induced central
CC nervous system disease.
CC -!- CATALYTIC ACTIVITY: L-arginine + NADPH + M O(2) = citrulline +
CC nitric oxide + NADP(+).
CC -!- COFACTOR: Heme. Binds one mole each of FAD and FMN. Also requires
CC tetrahydrobiopterin (BH4) which may stabilize the dimeric form of
CC the enzyme (By similarity).
CC -!- ENZYME REGULATION: Not stimulated by calcium/calmodulin (By

```

```

CC similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- INDUCTION: By lipopolysaccharide (LPS).
CC -!- SIMILARITY: Belongs to the NOS family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U31907; AAC39525.1; -.
CC HSP; P35228; 4NOS.
CC InterPro: IPR004030; NO synthase.
CC Pfam; PF02898; NO_synthase; 1.
CC PROSITE; PS60001; NOS; PARTIAL.
CC Oxidoreductase; NADP; FAD; FMN; Calmodulin-binding; Heme;
CC Multigene family.
FT NON_TER 1 1
FT NON_TER 162 162
FT SEQUENCE 162 AA; 17813 MW; 0457464D32287065 CRC64;
Query Match 2.8%; Score 6; DB 1; Length 162;
Best Local Similarity 100.0%; Pred.No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 LSLEIS 147
DB 129 LSLEIS 134

RESULT 78
FMCI ECOLI
ID FMCI ECOLI STANDARD; PRT; 170 AA.
AC P02971;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CFA/I fimbrial subunit B precursor (Colonization factor antigen I
DE subunit B) (CFA/I pilin) (CFA/I antigen).
GN CFAB.
OS Escherichia coli.
OG Plasmid NTP513.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89173309; PubMed=2564374;
RA Karjalainen T.K., Evans D.G., So M., Lee C.-H.;
RT "Molecular cloning and nucleotide sequence of the colonization factor
RT antigen I gene of Escherichia coli.";
RL Infect. Immun. 57:1126-1130(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89330163; PubMed=2569152;
RA Hamers A.M., Pel H.J., Willshaw G.A., Kusters J.G.,
RA van der Zeijst B.A.M., Gastra W.;
RT "The nucleotide sequence of the first two genes of the CFA/I fimbrial
RT operon of human enterotoxigenic Escherichia coli.";
RL Microb. Pathog. 6:297-309(1989).
RN [3]
RP SEQUENCE OF 24-170 FROM N.A.
RC STRAIN=078.H11 / H10407;
RX MEDLINE=82235736; PubMed=6124420;
RA Klemm P.;
RT "Primary structure of the CFA/I fimbrial protein from human
RT enterotoxigenic Escherichia coli strains.";
RL Eur. J. Biochem. 124:339-348(1982).
RN [4]
RP SEQUENCE OF 24-42.

```

```

RC STRAIN=078:H11 / H10407;
RX MEDLINE=90036735; PubMed=2572583;
RA Hall R.H., Maneval D.R. Jr., Collins J.H., Theibert J.L.,
RA Levine M.M.;
RT Purification and analysis of colonization factor antigen I, coli
RT surface antigen I, and coli surface antigen 3 fimbriae from
RL enterotoxigenic Escherichia coli.";
RL J. Bacteriol. 171:6372-6374(1989);
CC -!- FUNCTION: Fimbriae (also called pili), polar filaments radiating
CC from the surface of the bacterium to a length of 0.5-1.5
CC micrometers and numbering 100-300 per cell, enable bacteria to
CC colonize the epithelium of specific host organs.
CC -!- SUBUNIT: CFA/I FIMBRIAE ARE RATHER RIGID, THREAD-LIKE FILAMENTS OF
CC 0.5-1 MICROMETER, WITH AN APPARENT AXIAL HOLE, AND A DIAMETER OF
CC 7 NANOMETERS. A SINGLE CFA/I FIMBRIA CONSISTS OF ABOUT 100
CC IDENTICAL PROTEIN SUBUNITS.
CC -!- INDUCTION: CFA/I FIMBRIAE ARE ONLY EXPRESSED IN THE PRESENCE OF
CC THE POSITIVE REGULATOR CPAD.
CC -!- SIMILARITY: TO THE CSI FIMBRIAL SUBUNIT A (CSOA).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M55661; AAC41415.1; -.
DR PIR; A30589; YOECC1.
DR InterPro; IPR007540; Fimbril CSI.
DR Pfam; PF04449; Fimbril CSI; I.
DR Fimbril; Antigen; Plasmid; Signal.
FT SIGNAL 1 23
FT CHAIN 24 170 CFA/I FIMBRIAL SUBUNIT B.
FT CONFLICT 37 37 V -> A (IN REF. 2).
FT CONFLICT 76 76 D -> N (IN REF. 3).
FT CONFLICT 97 97 S -> A (IN REF. 3).
SQ SEQUENCE 170 AA; 17461 MW; 924347F07F33CDAB CRC64;

Query Match 2.8%; Score 6; DB 1; Length 170;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 FVAVSA 17
Db 16 FVAVSA 21
|||||

RESULT 79
APT ANASP STANDARD; PRT; 172 AA.
AC QVYNI3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adenine phosphoribosyltransferase (EC 2.4.2.7) (APRT).
GN APT OR ALR4582.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001).
CC -!- FUNCTION: Catalyzes a salvage reaction resulting in the formation

```

```

CC of AMP, that is energetically less costly than de novo synthesis.
CC -!- CATALYTIC ACTIVITY: AMP + diphosphate = adenine + 5-phospho-
CC alpha-D-ribose 1-diphosphate.
CC -!- PATHWAY: Purine salvage.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the purine/pyrimidine
CC phosphoribosyltransferase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP003596; BAB76281.1; -.
DR PIR; AF2378; AF2378.
DR HAMAP; MF_00004; -.
DR InterPro; IPR005764; Ade_phospho_transf.
DR InterPro; IPR002375; Pr/py_rp_transf.
DR InterPro; IPR000836; PRtransferase.
DR Pfam; PF00156; P-ribosyltran; 1.
DR TIGFAMs; TIGR01090; apt; 1.
DR PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.
DR TRANSFERase; Glycosyltransferase; Purine salvage; Complete proteome.
KW TRANSFERase; Glycosyltransferase; Purine salvage; Complete proteome.
SQ SEQUENCE 172 AA; 19144 MW; EA77762454D21A85 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 172;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 LAYKLG 123
Db 68 LAYKLG 73
|||||

RESULT 80
YA68_METH
ID YA68_METH STANDARD; PRT; 174 AA.
AC Q27140;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MTH1068.
GN MTH1068.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=96037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Tiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT DeltaH: functional analysis and comparative genomics."
RL J. Bacteriol. 179:7135-7155(1997).
CC -!- SIMILARITY: TO M.JANNASCHII MJ1559 AND A.FULGIDUS AF0814.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000877; AAB85557.1; -.
DR PIR; G69008; G69008.
DR InterPro; IPR004948; DUF265.
DR Pfam; PF03266; DUF265; 1.
KW Hypothetical protein; ATP-binding; Complete proteome.
FT NP_BIND 7 14 ATP (POTENTIAL).
SQ SEQUENCE 174 AA; 19425 MW; 646AAC1FED9AA2AC CRC64;

Query Match 2.8%; Score 6; DB 1; Length 174;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LLIAAV 10
DB 127 LLIAAV 132

RESULT 81
PHEB_AGLNE
ID PHEB_AGLNE STANDARD; PRT; 176 AA.
AC P28560;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE R-phycocerythrin beta chain.
GN CPEB OR R2EB.
OS Aglaethamion neglectum (Red alga).
OG Chloroplast.
OC Eukaryota; Rhodophyta; Florideophyceae; Ceramiales; Ceramiales;
OC Aglaethamion.
OX NCBI_TaxID=2765;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93144689; PubMed=7678762;
RA Apt K.B., Grossman A.R.;
RT "Characterization and transcript analysis of the major
RT phycobiliprotein subunit genes from Aglaethamion neglectum
RT (Rhodophyta).";
RL Plant Mol. Biol. 21:27-38(1993).
CC -!- FUNCTION: Light-harvesting photosynthetic bile pigment-protein
CC from the phycobiliprotein complex.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain.
CC -!- SUBCELLULAR LOCATION: Periphery of the rods of the phycobilisome.
CC -!- PTM: Contains three covalently linked bilin chromophores.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z11907; CA477962.1; -.
DR PIR; S30941; S30941.
DR HSP; O36004; I88D.
DR InterPro; IPR001659; Phycobilisome.
DR Pfam; PF00502; Phycobilisome; 1.
DR ProDom; PD000340; Phycobilisome; 1.
KW Phycobilisome; Electron transport; Photosynthesis; Bile pigment;
KW Chloroplast; Methylation.
FT MOD_RES 72 72 METHYLATION (BY SIMILARITY).
FT BINDING 50 50 PHYCOERYTHROBILIN CHROMOPHORE 1.
FT BINDING 61 61 PHYCOERYTHROBILIN CHROMOPHORE 1.
FT BINDING 82 82 PHYCOERYTHROBILIN CHROMOPHORE 2.
FT BINDING 158 158 PHYCOERYTHROBILIN CHROMOPHORE 3.
SQ SEQUENCE 176 AA; 18380 MW; 59E27FD5935DF6C3 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 176;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 8 AAVAFV 13
DB 137 AAVAFV 142

RESULT 82
Y045_CAUCR
ID Y045_CAUCR STANDARD; PRT; 178 AA.
AC Q9AC22;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0090 protein CC0045.
GN CC0045.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -!- SIMILARITY: Belongs to the UPF0090 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005679; BAK22033.1; AUT_INIT.
DR TIGR; E87254; E87254.
DR TIGR; CC0045; -.
DR HAMAP; MF_01077; -.
DR InterPro; IPR003728; DUF150.
DR Pfam; PF02576; DUF150; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 178 AA; 19706 MW; 709B182C6CA071A3 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 178;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 FKGEA 72
DB 111 FKGEA 116

RESULT 83
APT2_YEAST
ID APT2_YEAST STANDARD; PRT; 181 AA.
AC P36973; Q04090;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Adenine phosphoribosyltransferase 2 (EC 2.4.2.7) (APT 2).
GN APT2 OR YDR414C OR D9461.27.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
```

[1] SEQUENCE FROM N.A.
RP MEDLINE=95028163; PubMed=7941749;
RA Yuryev A., Corden J.L.;
RT "A Saccharomyces cerevisiae gene encoding a potential adenine
phosphoribosyltransferase.";
RL Yeast 10:659-662(1994).
RN [2]

SEQUENCE FROM N.A.
RP MEDLINE=97313263; PubMed=9169867;
RA Jacq C., Alt-Moerbe J., Andre B., Bateau N., Bloeker H., Blugeon C.,
Bargues M., Baron L., Becker A., Biteau N., Bloeker H., Blugeon C.,
Boskovic J., Brandt P., Bruckner M., Butrago M.J., Coster F.,
Delaveau T., Del Rey F., Dujon B., Eide L.G., Garcia-Cantalejo J.M.,
Goffeau A., Gomez-Peris A., Granotier C., Hanemann V., Hankeln T.,
Hohelsel J.D., Jaeger W., Jimenez A., Jonniaux J.-L., Kraemer C.,
Kuesner H., Laamanen P., Legros Y., Louis E.J., Moeller-Rieker S.,
Monnet A., Moro M., Mueller-Auer S., Nussbaumer B., Paricio N.,
Paulin L., Perea J., Perez-Alonso M., Perez-Ortin J.E., Pohl T.M.,
Prydz H., Purnelle B., Rasmussen S.W., Remacha M., Revuelta J.L.,
Rieger M., Salom D., Saluz H.P., Salz J.E., Saren A.-M., Schaefer M.,
Schaefer M., Schmidt E.R., Schneider C., Scholler P., Schwarz S.,
Urrutazu L.A., Verhaesselt P., Viessers S., Voet M., Volckaert G.,
Wagner G., Wambutt R., Wedler H., Wedler H., Woelfl S., Harris D.E.,
Bowman S., Brown D., Churcher C.M., Connor R., Dedman K., Gentles S.,
Hamlin N., Hunt S., Jones L., McDonald S., Murphy L., Niblett D.,
Odell C., Oliver K., Rajandream M.A., Richards C., Shorel L.,
Walsh S.V., Barrell B.G., Dietrich F.S., Mulligan J.T., Allen E.,
Arayo R., Aviles E., Berno A., Carpenter J., Chen E., Cherry J.M.,
Chung E., Duncan M., Hunkle-Smith S., Hyman R.W., Komp C.,
Lashkari D., Lew H., Lin D., Mosedale D., Nakahara K., Namath A.,
Oefner P., Oh C., Petel F.X., Roberts D., Schramm S., Schroeder M.,
Shogren T., Shroff N., Winant A., Yelton M.A., Botstein D.,
Davis R.W., Johnston M., Andrews S., Brinkman R., Cooper J., Ding H.,
Du Z., Favell A., Fulton L., Gattung S., Greco T., Hallsworth K.,
Hawkins J., Hillier L.W., Jier M., Johnson D., Johnston L.,
Kirsten J., Kucaba T., Langston Y., Latreille P., Le T., Mardis E.,
Menezes S., Miller N., Nhan M., Pauley A., Peluso D., Rifkin L.,
Riles L., Taich A., Trevasik E., Vignati D., Wilcox L., Wohldman P.,
Vaudin M., Wilson R., Waterston R., Alberman K., Hani J., Heumann K.,
Kleine K., Mewes H.-W., Zollner A., Zaccaria P.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IV.";
RL Nature 387:75-78(1997).

CC -!- FUNCTION: Catalyzes a salvage reaction resulting in the formation
of AMP, that is energetically less costly than de novo synthesis.
CC -!- CATALYTIC ACTIVITY: AMP + diphosphate = adenine + 5-phospho-
alpha-D-ribose 1-diphosphate.
CC -!- PATHWAY: Purine salvage.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the purine/pyrimidine
phosphoribosyltransferase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L14434; AAA62848.1; -;
DR EMBL; U33007; AAB64883.1; -;
DR PIR; S69721; S69721. -;
DR GERMOnline; 140933; -;
DR SGD; S0002849; APT2.
DR InterPro; IPR002375; Pr/py ip transf.
DR Pfam; PF00156; Pribosyltran; 1.
DR PROSITE; PS00103; PUR_PYR_PR_TRANSF; 1.
KW Transferase; Glycosyltransferase; Purine salvage; Multigene family.
FT CONFLICT 57 57 A -> G (IN REF. 2).
SQ SEQUENCE 181 AA; 20000 MW; 56AB4D41380DFEB3 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 181;
Best Local Similarity 100.0%; Pred.No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 76 IEARGL 81
DB 67 IEARGL 72

RESULT 84
DSBE SALT
ID_DSBE SALT STANDARD; PRT; 185 AA.
AC Q8AFES;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Thiol:disulfide interchange protein dsbE (Cytochrome c biogenesis
protein ccmG).
DE (DSBE1 OR CCMG1 OR STY2474 OR T0616) AND
GN (DSBE2 OR CCMG2 OR STY3965 OR T3705).
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
Feltwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
Kregh A., Larsen T.S., Leather S., Moulle S., O'Gaora P., Parry C.,
Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
CC -!- FUNCTION: Involved in disulfide bond formation. Catalyzes a late,
reductive step in the assembly of periplasmic c-type cytochromes,
probably the reduction of disulfide bonds of the apocytochrome c
to allow covalent linkage with the heme. Possible subunit of a
heme lyase (By similarity).
CC -!- SUBCELLULAR LOCATION: Mostly periplasmic; anchored in the inner
membrane (By similarity).
CC -!- SIMILARITY: Belongs to the thioredoxin family. DsbE subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL627274; CAD07480.1; -;
DR EMBL; AL627280; CAD03181.1; -;
DR EMBL; AE016836; AAO68321.1; -;
DR EMBL; AE016846; AAO71200.1; -;
DR InterPro; IPR004759; DsbE.
DR InterPro; IPR006662; Thiored.
DR ProDom; PD003679; DsbE; 1.

DR TIGRFAMS; TIGR00385; dsBE; 1.
 DR PROSITE; PS00194; THIOREDOXIN; 1.
 KW Cytochrome c-type biogenesis; Redox-active center; Transmembrane;
 KW Inner membrane; Complete proteome.
 FT DOMAIN 1 4 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 5 25 POTENTIAL.
 FT DOMAIN 26 185 PERIPLASMIC (POTENTIAL).
 FT DISULFID 80 83 REDOX-ACTIVE (BY SIMILARITY).
 SQ SEQUENCE 185 AA; 20691 MW; AFA2A848B5828E59 CRC64;
 Query Match 2.8%; Score 6; DB 1; Length 185;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 FLLIAA 9
 DB 13 FLLIAA 18
 RESULT 85
 DSBE_SALTY STANDARD; PRT; 185 AA.
 AC Q8XFK6;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Thiol-disulfide interchange protein dsBE (Cytochrome c biogenesis
 DE protein cmcg)
 GN (DSBE1 OR CCMG1 OR STM2248) AND (DSBE2 OR CCMG2 OR STM3813).
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2.";
 RL Nature 413:852-856(2001).
 CC -!- FUNCTION: Involved in disulfide bond formation. Catalyzes a late,
 CC reductive step in the assembly of periplasmic c-type cytochromes,
 CC probably the reduction of disulfide bonds of the apocytochrome c
 CC to allow covalent linkage with the heme. Possible subunit of a
 CC heme lyase (by similarity).
 CC -!- SUBCELLULAR LOCATION: Mostly periplasmic; anchored in the inner
 CC membrane (By similarity).
 CC -!- SIMILARITY: Belongs to the thioredoxin family. Dsbe subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE008877; AAL22672.1; -
 CC EMBL; AE008800; AAL21150.1; -
 CC STycene; SG????; dsBE1.
 CC STycene; SG????; dsBE2.
 CC InterPro; IPR004799; Dsbe.
 CC InterPro; IPR006662; ThioRed.
 CC InterPro; IPR006663; ThioRedox_dom2.
 CC ProDom; PD003679; Dsbe; 1.
 CC TIGRFAMS; TIGR00385; dsBE; 1.
 DR PROSITE; PS00194; THIOREDOXIN; 1.
 KW Cytochrome c-type biogenesis; Redox-active center; Transmembrane;
 KW Inner membrane; Complete proteome.
 FT DOMAIN 1 4 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 5 25 POTENTIAL.
 FT DOMAIN 26 185 PERIPLASMIC (POTENTIAL).
 FT DISULFID 80 83 REDOX-ACTIVE (BY SIMILARITY).
 SQ SEQUENCE 185 AA; 20691 MW; AFA2A848B5828E59 CRC64;
 Query Match 2.8%; Score 6; DB 1; Length 185;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 FLLIAA 9
 DB 13 FLLIAA 18
 RESULT 86
 DYSR_HUMAN STANDARD; PRT; 186 AA.
 AC P00374; Q14130;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Dihydrofolate reductase (EC 1.5.1.3).
 GN DHFR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84162075; PubMed=6323448;
 RA Chen M.-J., Shimada T., Moulton A.D., Cline A., Humphries R.K.,
 RA Maizel J., Nienhuis A.W.;
 RA "The functional human dihydrofolate reductase gene.";
 RL J. Biol. Chem. 259:3933-3943(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83183667; PubMed=6687716;
 RA Masters J.N., Attardi G.;
 RA "The nucleotide sequence of the cDNA coding for the human
 RA dihydrofolate acid reductase.";
 RL Gene 21:59-63(1983).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84267838; PubMed=6235374;
 RA Yang J.K., Masters J.N., Attardi G.;
 RA "Human dihydrofolate reductase gene organization. Extensive
 RA conservation of the G + C-rich 5' non-coding sequence and strong
 RA intron size divergence from homologous mammalian genes.";
 RL J. Mol. Biol. 176:169-187(1984).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Eve;
 RX MEDLINE=2388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length

KW Inner membrane; Complete proteome.
 FT DOMAIN 1 4 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 5 25 POTENTIAL.
 FT DOMAIN 26 185 PERIPLASMIC (POTENTIAL).
 FT DISULFID 80 83 REDOX-ACTIVE (BY SIMILARITY).
 SQ SEQUENCE 185 AA; 20679 MW; 9B731AA457929C68 CRC64;
 Query Match 2.8%; Score 6; DB 1; Length 185;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 FLLIAA 9
 DB 13 FLLIAA 18
 RESULT 86
 DYSR_HUMAN STANDARD; PRT; 186 AA.
 AC P00374; Q14130;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Dihydrofolate reductase (EC 1.5.1.3).
 GN DHFR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84162075; PubMed=6323448;
 RA Chen M.-J., Shimada T., Moulton A.D., Cline A., Humphries R.K.,
 RA Maizel J., Nienhuis A.W.;
 RA "The functional human dihydrofolate reductase gene.";
 RL J. Biol. Chem. 259:3933-3943(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83183667; PubMed=6687716;
 RA Masters J.N., Attardi G.;
 RA "The nucleotide sequence of the cDNA coding for the human
 RA dihydrofolate acid reductase.";
 RL Gene 21:59-63(1983).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84267838; PubMed=6235374;
 RA Yang J.K., Masters J.N., Attardi G.;
 RA "Human dihydrofolate reductase gene organization. Extensive
 RA conservation of the G + C-rich 5' non-coding sequence and strong
 RA intron size divergence from homologous mammalian genes.";
 RL J. Mol. Biol. 176:169-187(1984).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Eve;
 RX MEDLINE=2388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[5]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=88254806; PubMed=3383852;
RA Cefner C., D'Arcy A., Winkler P.K.;
RT "Crystal structure of human dihydrofolate reductase complexed with
RL folate.";
RL Eur. J. Biochem. 174:377-385(1988).
[6]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=91064350; PubMed=2248959;
RA Davies J.F., Delcamp T.J., Prendergast N.J., Ashford V.A.,
RA Freisheim J.H., Kraut J.;
RT "Crystal structures of recombinant human dihydrofolate reductase
RL complexed with folate and 5-deazafoate.";
RL Biochemistry 29:9467-9479(1990).
[7]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=98042281; PubMed=9374868;
RA Cody V., Galitsky N., Luft J.R., Pangborn W., Blakley R.L.,
RA Gangjee A.;
RT "Comparison of two independent crystal structures of human
RL dihydrofolate reductase ternary complexes reduced with nicotinamide
RT adenine dinucleotide phosphate and the very tight-binding inhibitor
RT PT523.";
RL Biochemistry 36:13897-13903(1997).
[8]
RP STRUCTURE BY NMR.
RX MEDLINE=92118795; PubMed=1731871;
RA Stockman B.J., Nirmala N.R., Wagner G., Delcamp T.J., Deyarman M.T.,
RA Freisheim J.H.;
RT "Sequence-specific 1H and 15N resonance assignments for human
RL dihydrofolate reductase in solution.";
RL Biochemistry 31:218-229(1992).
CC -1- CATALYTIC ACTIVITY: 5,6,7,8-tetrahydrofolate + NADP(+) = 7,8-
CC dihydrofolate + NADPH.
CC -1- PATHWAY: Essential step for de novo glycine and purine synthesis,
CC DNA precursor synthesis, and for the conversion of dUMP to dTMP.
CC -1- SIMILARITY: Belongs to the dihydrofolate reductase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00140; AAA58485.1; -;
DR EMBL; V00507; CAA23765.1; -;
DR EMBL; J00139; AAA58484.1; -;
DR EMBL; K01612; AAA58484.1; JOINED.
DR EMBL; K01613; AAA58484.1; JOINED.
DR EMBL; J00138; AAA58484.1; JOINED.
DR EMBL; K01614; AAA58484.1; JOINED.
DR EMBL; X00855; CAA25409.1; -;
DR EMBL; X00856; CAA25409.1; JOINED.
DR EMBL; X00857; CAA25409.1; JOINED.
DR EMBL; X00858; CAA25409.1; JOINED.
DR EMBL; X00859; CAA25409.1; JOINED.
DR EMBL; BC000192; AAH00192.1; -;
DR EMBL; BC003584; AAH03584.2; -;
DR PIR; A22551; RDHUF.
DR PDB; 1DHF; 15-JUL-92.
DR PDB; 2DHF; 15-JAN-93.
DR PDB; 1DRF; 15-JUL-92.
DR PDB; 1DLR; 20-APR-95.
DR PDB; 1DLS; 20-APR-95.
DR PDB; 1HFP; 28-JAN-98.
DR PDB; 1HFR; 28-JAN-98.
DR PDB; 1OHJ; 29-APR-98.

DR PDB; 1OHK; 27-MAY-98.
DR PDB; 1BOZ; 12-AUG-98.
DR PDB; 1KMS; 10-JUL-02.
DR PDB; 1KVV; 10-JUL-02.
DR HSC-2DPAGE; P00374; HUMAN.
DR Genew; HGNC:2861; DHFR.
DR MIM; 126060;
DR GO; GO:0004146; F:dihydrofolate reductase activity; NAS.
DR GO; GO:0006545; P:glycine biosynthesis; NAS.
DR GO; GO:0009165; P:nucleotide biosynthesis; NAS.
DR InterPro; IPR001796; DHFR.
DR Pfam; PF00186; Dihfolate_red; 1.
DR PRINTS; PR00070; DHFR.
DR PROSITE; PS00075; DHFR; 1.
KW Oxidoreductase; NADP; One-carbon metabolism; 3D-structure.
FT INIT_MET 0
FT STRAND 4 10
FT TURN 12 13
FT STRAND 15 17
FT TURN 15 17
FT HELIX 28 39
FT TURN 43 44
FT STRAND 47 53
FT HELIX 54 59
FT HELIX 62 64
FT TURN 68 69
FT STRAND 71 75
FT TURN 84 85
FT STRAND 88 90
FT HELIX 93 99
FT TURN 100 101
FT HELIX 103 106
FT TURN 107 108
FT STRAND 109 114
FT HELIX 118 126
FT STRAND 130 138
FT TURN 146 147
FT TURN 153 155
FT STRAND 157 158
FT STRAND 162 162
FT TURN 163 164
FT STRAND 165 165
FT STRAND 170 172
FT TURN 173 174
FT STRAND 175 186
SQ SEQUENCE 186 AA; 21321 MW; EBB9E6A3ECAC8CEDB CRC64;
Query Match 2.8%; Score 6; DB 1; Length 186;
Best Local Similarity 100.0%; Pred.No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 180 GVLSDV 185
Db 164 GVLSDV 169
|||||
RESULT 87
DENR_CAEEL
ID DENR_CAEEL STANDARD; PRT; 192 AA.
AC Q9NAH4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Density-regulated protein homolog.
GN Y47D3A.21.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Matthews L.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

```
CC -!- SIMILARITY: Contains 1 SU11 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL117202; CAB55081.1; -.
DR PIR; T31532; T31532.
DR WormPep; Y47D3A.21; CE22051.
DR InterPro; IPR005873; TrpL.
DR InterPro; IPR001950; TrpL SU11.
DR Pfam; PF01253; SU11; 1.
DR TIGRFAMs; TIGR01159; DRP1; 1.
DR PROSITE; PS50296; SU11; 1.
DR Hypothetical protein.
KW Hypothetical protein.
FT DOMAIN 117..176
FT SEQUENCE 192 AA; 20349 MW; E0C12759A38A5BA6 CRC64;
SU11.
SU11.
Query Match 2.8%; Score 6; DB 1; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 144 LEISDE 149
DB 63 LEISDE 68
-----
RESULT 88
TSAA_BUCAI
ID TSAA_BUCAI STANDARD; PRT; 197 AA.
AC P57279;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable peroxiredoxin.
GN BUI82.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS.";
RL Nature 407:81-86(2000).
CC -!- SUPRACELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the alpC/TSA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP001118; BAB12899.1; -.
DR HSSP; P30041; 1PRX.
DR InterPro; IPR000866; AlpC-TSA.
DR Pfam; PF00578; AlpC-TSA; 1.
KW Antioxidant; Complete proteome.
FT ACT SITE 50..50
FT SEQUENCE 197 AA; 22363 MW; D08C2CA66C256376 CRC64;
BY SIMILARITY.
Query Match 2.8%; Score 6; DB 1; Length 197;
```

```
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 93 KGEEGI 98
DB 177 KGEEGI 182
-----
RESULT 89
Y5E0_VIBPA
ID Y5E0_VIBPA STANDARD; PRT; 199 AA.
AC Q87J78;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0217 protein VPA0140.
GN VPA0140.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Tijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
CC -!- SIMILARITY: Belongs to the UPF0217 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP005084; BAC61483.1; -.
CC HAMAP; MF_00587; -.
CC InterPro; IPR007158; DUF358.
CC Pfam; PF04013; DUF358; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 199 AA; 22277 MW; DE51A3F190DE2F4F CRC64;
Query Match 2.8%; Score 6; DB 1; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 LIAAVA 11
DB 83 LIAAVA 88
-----
RESULT 90
LEUD_EC057
ID LEUD_EC057 STANDARD; PRT; 200 AA.
AC Q8XA01;
DT 28-FEB-2003 (Rel. 41, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 3-isopropylmalate dehydratase small subunit (EC 4.2.1.33)
DE (isopropylmalate isomerase) (Alpha-IPM isomerase) (IPMI).
GN LEUD OR Z0080 OR ECS0075.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
```



```
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11205551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postai G., Hackett J., Klank S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamocis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kikuta S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [3]
RP SEQUENCE OF 1-4.
RC STRAIN=K12 / MG1655;
RX MEDLINE=927426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE OF 1-4.
RC STRAIN=K12 / W3110;
RX MEDLINE=98263447; PubMed=9600841;
RA Wilkins M.R., Gasteliger E., Tonella L., Ou K., Tyler M.,
RA Sanchez J.-C., Gooley A.A., Walsh B.J., Sairoch A., Appel R.D.,
RA Williams K.L., Hochstrasser D.F.;
RT "Protein identification with N and C-terminal sequence tags in
proteome projects.";
RL J. Mol. Biol. 278:599-608(1998).
CC -!- FUNCTION: Catalyzes the isomerization between 2-isopropylmalate
and 3-isopropylmalate, via the formation of 2-isopropylmaleate +
H(2)O.
CC -!- CATALYTIC ACTIVITY: 3-isopropylmalate = 2-isopropylmaleate +
H(2)O.
CC -!- CATALYTIC ACTIVITY: 2-isopropylmaleate + H(2)O = 2-
isopropylmalate.
CC -!- PATHWAY: Leucine biosynthesis; second step.
CC -!- SUBUNIT: Heterodimer of leuc and leud (By similarity).
CC -!- SIMILARITY: Belongs to the leud family. Leud 1 subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005184; AAC54375.1; -.
DR EMBL; AP002550; BAB33498.1; -.
DR PIR; C85489; C85489.
DR PIR; C90638; C90638.
DR HANAP; MF_01031; -.
DR InterPro; IPR000573; Aconitase_C.
DR Pfam; PF00694; Aconitase_C; 1.
KW Leucine biosynthesis; Lyase; Complete proteome.
FT INIT MET 0
FT BY SIMILARITY.
SQ SEQUENCE 200 AA; 22388 MW; E79D6CA3E66AB1D9 CRC64;
Query Match 2.8%; Score 6; DB 1; Length 200;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 34 DDATIAA 39
DB 185 DDATIAA 190
RESULT 91
LEUD_ECOLI STANDARD; PRT; 200 AA.
AC P30126;
DT 01-APR-1993 (Rel. 25, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 3-isopropylmalate dehydratase small subunit (EC 4.2.1.33)
DE (isopropylmalate isomerase) (Alpha-IPM isomerase) (IPMI).
GN LEUD OR B0071.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
```

```
RN SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92334977; PubMed=1630901;
RA Yura T., Mori H., Nagai H., Negata T., Ishihama A., Fujita N.,
RA Isono K., Mizobuchi K., Nakata A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
the 0-2.4 min region.";
RL Nucleic Acids Res. 20:3305-3308(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE OF 1-4.
RC STRAIN=K12 / W3110;
RX MEDLINE=98263447; PubMed=9600841;
RA Wilkins M.R., Gasteliger E., Tonella L., Ou K., Tyler M.,
RA Sanchez J.-C., Gooley A.A., Walsh B.J., Sairoch A., Appel R.D.,
RA Williams K.L., Hochstrasser D.F.;
RT "Protein identification with N and C-terminal sequence tags in
proteome projects.";
RL J. Mol. Biol. 278:599-608(1998).
CC -!- FUNCTION: Catalyzes the isomerization between 2-isopropylmalate
and 3-isopropylmalate, via the formation of 2-isopropylmaleate +
H(2)O.
CC -!- CATALYTIC ACTIVITY: 3-isopropylmalate = 2-isopropylmaleate +
H(2)O.
CC -!- CATALYTIC ACTIVITY: 2-isopropylmaleate + H(2)O = 2-
isopropylmalate.
CC -!- PATHWAY: Leucine biosynthesis; second step.
CC -!- SUBUNIT: Heterodimer of leuc and leud (By similarity).
CC -!- SIMILARITY: Belongs to the leud family. Leud 1 subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10483; BAB96640.1; -.
DR EMBL; AE000117; AAC73182.1; -.
DR PIR; S40585; S40585.
DR SWISS-2DPAGE; P30126; COLI.
DR Ecogene; EG11575; leud.
DR HAMAP; MF_01031; -.
DR InterPro; IPR000573; Aconitase_C.
DR Pfam; PF00694; Aconitase_C; 1.
KW Leucine biosynthesis; Lyase; Complete proteome.
FT INIT MET 0
FT BY SIMILARITY.
SQ SEQUENCE 200 AA; 22356 MW; E54E3A3EBE697ED8 CRC64;
Query Match 2.8%; Score 6; DB 1; Length 200;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 34 DDATIAA 39
DB 185 DDATIAA 190
RESULT 92
ENO_CAMFEE STANDARD; PRT; 207 AA.
ID ENO_CAMFEE
AC Q30885;
DT 15-JUL-1999 (Rel. 38, Created)
```

DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
 DE glycerate hydro-lyase) (Fragment).
 GN ENO.
 OS Campylobacter fetus.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Campylobacteraceae; Campylobacter.
 OX NCBI_TaxID=196;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=230;
 RX MEDLINE=98053968; PubMed=9393719;
 RA Workin J., Shedd O.L., Blaser M.J.;
 RT "Nested DNA inversion of Campylobacter fetus S-layer genes is recA
 RT dependent".
 RL J. Bacteriol. 179:7523-7529(1997).
 CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
 CC H₂O.
 CC -1- COFACTOR: Magnesium is required for catalysis and for stabilizing
 CC the dimer (By similarity).
 CC -1- PATHWAY: Glycolysis.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the enolase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF020677; AAB86924.1; -.
 DR HSSP; P56252; 1PDZ.
 DR HAMAP; MF 00318; -; 1.
 DR InterPro; IPR000941; Enolase.
 DR Pfam; PF00113; enolase; 1.
 DR PRINTS; PR00148; ENOLASE.
 DR ProDom; PD000902; Enolase; 1.
 DR ProSITE; PS00164; ENOLASE; PARTIAL.
 KW Lyase; Glycolysis; Magnesium.
 FT ACT SITE 154
 FT NON-TER 207 207 BY SIMILARITY.
 SQ SEQUENCE 207 AA; 21833 MW; B6752AC9624EB393 CRC64;
 Query Match 2.8%; Score 6; DB 1; Length 207;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DR EMBL; AF020677; AAB86924.1; -.
 DR HSSP; P56252; 1PDZ.
 DR HAMAP; MF 00318; -; 1.
 DR InterPro; IPR000941; Enolase.
 DR Pfam; PF00113; enolase; 1.
 DR PRINTS; PR00148; ENOLASE.
 DR ProDom; PD000902; Enolase; 1.
 DR ProSITE; PS00164; ENOLASE; PARTIAL.
 KW Lyase; Glycolysis; Magnesium.
 FT ACT SITE 154
 FT NON-TER 207 207 BY SIMILARITY.
 SQ SEQUENCE 207 AA; 21833 MW; B6752AC9624EB393 CRC64;
 Query Match 2.8%; Score 6; DB 1; Length 207;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 31 KAIDDA 36
 Db 87 KAIDDA 92
 RESULT 93
 EFH2 TRYCR STANDARD; PRT; 208 AA.
 AC F41049;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE EF-hand protein 5 variant 2 (EFH5) (Calmodulin-ubiquitin associated
 DE protein CUB2.8).
 OS Trypanosoma cruzi.
 OC Eukaryota; Eulenczoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5693;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93149197; PubMed=8381204;
 RA Ajioke J., Swindle J.T.;

RT "The calmodulin-ubiquitin associated genes of Trypanosoma cruzi:
 RT their identification and transcription.";
 RL Mol. Biochem. Parasitol. 57:127-136(1993).
 CC -1- DOMAIN: This protein has four EF-hand domains, two of which may be
 CC functional calcium-binding sites.
 CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS. BELONGS TO
 CC THE EFH5 SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; L01584; AAA30172.1; -.
 DR HSSP; P02593; 1CDM.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 1.
 DR ProDom; PD000012; EF-hand; 1.
 DR PROSITE; PS00018; EF HAND; FALSE_NEG.
 KW Calcium-binding; Repeat.
 FT DOMAIN 72 84 ANCESTRAL CALCIUM SITE 1 (POTENTIAL).
 FT CA_BIND 112 123 EF-HAND 2 (POTENTIAL).
 FT CA_BIND 148 159 EF-HAND 3 (POTENTIAL).
 FT DOMAIN 184 195 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
 SQ SEQUENCE 208 AA; 23671 MW; 15516948A9567E3 CRC64;
 Query Match 2.8%; Score 6; DB 1; Length 208;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 87 QGDANV 92
 Db 10 QGDANV 15
 RESULT 94
 UPF_OCEIH STANDARD; PRT; 209 AA.
 ID UPF_OCEIH STANDARD; PRT; 209 AA.
 AC Q8EM74;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Uracil phosphoribosyltransferase (EC 2.4.2.5) (UMP pyrophosphorylase)
 DE (UPRTase).
 DE UPF OR 0B2984.
 OS Oceanobacillus iheyensis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
 OX NCBI_TaxID=182710;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
 RX MEDLINE=22220767; PubMed=12235376;
 RA Takami H., Takaki F., Uchiyama I.;
 RT "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya
 RT Ridge and its unexpected adaptive capabilities to extreme
 RT environments".
 RL Nucleic Acids Res. 30:3927-3935(2002).
 CC -1- CATALYTIC ACTIVITY: UMP + diphosphate = uracil + 5-phospho-alpha-
 CC D-ribose 1-diphosphate.
 CC -1- COFACTOR: Magnesium (By similarity).
 CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- SIMILARITY: Belongs to the UPRTase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

```
CC EMBL; AP004603; BAC14940.1; -.
DR HAVAP; MF 01218; -.
DR InterPro; IPR000836; PRTransferase.
DR InterPro; IPR005765; Ura_phospho_trans.
DR Pfam; PF00156; Pribosyltran; 1.
DR TIGRFAMs; TIGR01091; upp; 1.
KW Transferase; Glycosyltransferase; Magnesium; Complete proteome.
FT DOMAIN 127 139 PRPP-BINDING (BY SIMILARITY).
FT DOMAIN 193 201 UMP-BINDING, URACIL PART (BY SIMILARITY).
SQ SEQUENCE 209 AA; 22914 MW; 54E6FF936976509A CRC64;

Query Match 2.8%; Score 6; DB 1; Length 209;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 DATAAI 40
DB 142 DATAAI 147

RESULT 95
UPP_RHIME STANDARD; PRT; 209 AA.
AC Q92T49;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uracil phosphoribosyltransferase (EC 2.4.2.9) (UMP pyrophosphorylase)
DE (UPRTase).
GN UPP OR R00131 OR SMC04121.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales.
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler P., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetalle D., Puehler A., Purnelle B., Ransperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -1- CATALYTIC ACTIVITY: UMP + diphosphate = uracil + 5-phospho-
CC alpha-D-ribose 1-diphosphate.
CC -1- COFACTOR: Magnesium (by similarity).
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SIMILARITY: Belongs to the UPRTase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; AL591782; CAC041518.1; -.
DR HAVAP; MF 01218; -.
DR InterPro; IPR000836; PRTransferase.
DR InterPro; IPR005765; Ura_phospho_trans.
DR Pfam; PF00156; Pribosyltran; 1.
DR TIGRFAMs; TIGR01091; upp; 1.
KW Transferase; Glycosyltransferase; Magnesium; Complete proteome.
FT DOMAIN 127 139 PRPP-BINDING (BY SIMILARITY).
FT DOMAIN 193 201 UMP-BINDING, URACIL PART (BY SIMILARITY).
SQ SEQUENCE 209 AA; 23032 MW; 50386F36DF642DD7 CRC64;

Query Match 2.8%; Score 6; DB 1; Length 209;
```

```
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 AIAAIE 41
DB 140 AIAAIE 145

RESULT 96
YPJC_BACSU STANDARD; PRT; 215 AA.
AC P42978;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein YPJC.
GN YPJC OR JOUC OR BSU22510.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / Marburg;
RX MEDLINE=96349105; PubMed=8760912;
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
RA Serror P.;
RT "Sequence analysis of the Bacillus subtilis chromosome region between
RT the serA and kds loci cloned in a yeast artificial chromosome.";
RL Microbiology 142:2005-2016(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Azevedo V., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Hoisappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapinus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche B., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Taccani E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RT Nature 390:249-256(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: TO B.SUBTILIS YITT AND YQFU.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)
```

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L38424; AAA92872.1; -.
DR EMBL; L47709; AAB38440.1; -.
DR EMBL; Z89115; CAB14167.1; -.
DR FIR; C69937; C69937.
DR Subtilist; BG11209; ypjC.
DR InterPro; IPR003740; DUF161.
DR Pfam; PF02588; DUF161; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 31 51 POTENTIAL.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 104 124 POTENTIAL.
SQ SEQUENCE 215 AA; 23582 MW; D314CF7225F8A983 CRC64;

Query Match      2.8%; Score 6; DB 1; Length 215;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 AFVAVS 16
DB 189 AFVAVS 194

RESULT 97
PRO2_XANCG
ID PRO2_XANCG STANDARD; PRT; 216 AA.
AC P37828;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-FEB-1996 (Rel. 33; Last sequence update)
DT 01-FEB-1996 (Rel. 33; Last annotation update)
DE Pathogenicity-related ORF2.
OS Xanthomonas campestris (pv. glycines).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=36865;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92193281; PubMed=1312532;
RA Hwang I., Lim S.M., Shaw P.D.;
RT Cloning and characterization of pathogenicity genes from Xanthomonas campestris pv. glycines.
RL J. Bacteriol. 174:1923-1931 (1992).
CC -!- FUNCTION: IMPORTANT FOR PATHOGENICITY.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE FLIP/MOPC/SPAP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M64094; AAA69368.1; -.
DR InterPro; IPR005838; TypeIII_P.
DR Pfam; PF00813; Flp; 1.
DR PRINTS; PR01302; TYPE3IMPPROT.
DR ProDom; PD002586; TypeIII_P; 1.
DR PROSITE; PS01060; FLIP_1; 1.
DR PROSITE; PS01061; FLIP_2; 1.
DR Virulence; Transmembrane; Flagellum.
FT TRANSMEM 6 26 POTENTIAL.
FT TRANSMEM 55 75 POTENTIAL.
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 193 213 POTENTIAL.
SQ SEQUENCE 216 AA; 23660 MW; 931E4A0400E6CF6C CRC64;

Query Match      2.8%; Score 6; DB 1; Length 216;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 AFVAVS 16
DB 189 AFVAVS 194

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 201 VLAPAF 206
DB 141 VLAPAF 146

RESULT 98
HRCR_RALSO
ID HRCR_RALSO STANDARD; PRT; 217 AA.
AC C52488;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Hypersensitivity response secretion protein hrcr.
GN HRCR OR HRP1 OR RSP0860 OR RS01631.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OG Plasmid megaplasmid.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GMI1000;
RX MEDLINE=95349395; PubMed=7623665;
RA van Gijsegem F., Gough C.L., Zischek C., Niqueux E., Arlat M.,
RA Genin S., Barberis P., German S., Castello P., Boucher C.A.;
RT "The hrp gene locus of Pseudomonas solanacearum, which controls the production of a type III secretion system, encodes eight proteins related to components of the bacterial flagellar biogenesis complex."
RL Mol. Microbiol. 15:1095-1114 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=GMI1000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangerot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Signier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
RL Nature 415:497-502 (2002).
CC -!- FUNCTION: INVOLVED IN THE SECRETION OF POPA, A PROTEINACEOUS ELICITOR OF THE HYPERSENSITIVITY RESPONSE IN PLANTS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE FLIP/MOPC/SPAP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ245811; CAB58247.1; -.
DR EMBL; AL646081; CAD18011.1; -.
DR FIR; S61851; S61851.
DR InterPro; IPR005838; TypeIII_P.
DR Pfam; PF00813; Flp; 1.
DR PRINTS; PR01302; TYPE3IMPPROT.
DR ProDom; PD002586; TypeIII_P; 1.
DR TIGRFAMs; TIGR01102; yscR; 1.
DR PROSITE; PS01060; FLIP_1; 1.
DR PROSITE; PS01061; FLIP_2; 1.
DR Hypersensitivity response; Transmembrane; Plasmid; Complete proteome.
FT TRANSMEM 6 26 POTENTIAL.
FT TRANSMEM 52 72 POTENTIAL.
FT TRANSMEM 158 178 POTENTIAL.
FT TRANSMEM 190 210 POTENTIAL.
SQ SEQUENCE 217 AA; 23739 MW; 79745499E2E2D80A CRC64;
```

```
Query Match      2.8%; Score 6; DB 1; Length 217;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 201 VLAPAF 206
DB 142 VLAPAF 147

RESULT 99
YQIY_BACSU STANDARD; PRT; 219 AA.
AC PS4536;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable amino-acid ABC transporter permease protein yqiy.
GN YQIY OR BSU3970.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=97124195; PubMed=8969508;
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M., Kobayashi Y.;
RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of the Bacillus subtilis genome containing the skin element and many sporulation genes.";
RL Microbiology 142:3103-3111(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N., Gim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Guisepi G., Guy B.J., Hega K., Haeleth J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.P., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C., Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogasawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B., Sokoluk A., Tacconi E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tanakashi A., Tanaka T., Terpestra P., Tognoni A., Toato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H., Weitzengger T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus subtilis";
RL Nature 390:249-256(1997).
CC -!- FUNCTION: Probably part of a binding-protein-dependent transport system for an amino acid. Probably responsible for the translocation of the substrate across the membrane.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the binding-protein-dependent transport system permease family. HisMQ subfamily.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
EMBL; D84432; BAA12605.1; -
EMBL; Z99116; CAS14328.1; -
PIR; G69962; G69962.
DR Subtilisin; BG11728; yqiy.
DR InterPro; IPR00515; BPD transp.
DR Pfam; PF00528; BPD transp; 1.
DR PROSITE; PS50928; ABC_TM1; 1.
KW Hypothetical protein; Transport; Amino-acid transport; Transmembrane; Complete proteome.
FT TRANSMEM 19 39 POTENTIAL.
FT TRANSMEM 66 86 POTENTIAL.
FT TRANSMEM 88 108 POTENTIAL.
FT TRANSMEM 149 169 POTENTIAL.
FT TRANSMEM 187 207 POTENTIAL.
SQ SEQUENCE 219 AA; 23897 MW; 041949734A5E19E5 CRC64;

Query Match      2.8%; Score 6; DB 1; Length 219;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVALSL 144
DB 93 VVALSL 98

RESULT 100
NHB2_RHOSH STANDARD; PRT; 226 AA.
AC P29379;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Low-molecular weight cobalt-containing nitrile hydratase subunit beta (EC 4.2.1.84) (L-nitrilase) (L-NHase).
OS Rhodococcus rhodochrous.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.
OX NCBI_TaxID=1829;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=J1;
RX MEDLINE=92096459; PubMed=1840499;
RA Kobayashi M., Nishiyama M., Nagasawa T., Horinouchi S., Beppu T., Yamada H.;
RT "Cloning, nucleotide sequence and expression in Escherichia coli of two cobalt-containing nitrile hydratase genes from Rhodococcus rhodochrous J1.";
RL Biochim. Biophys. Acta 1129:23-33(1991).
CC -!- FUNCTION: NHASE CATALYZES THE HYDRATION OF VARIOUS NITRILE COMPOUNDS TO THE CORRESPONDING AMIDES. INDUSTRIAL PRODUCTION OF ACRYLAMIDE IS NOW BEING DEVELOPED USING SOME OF THESE ENZYMES.
CC -!- CATALYTIC ACTIVITY: An aliphatic amide = a nitrile + H(2)O.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain.
CC -!- INDUCTION: By cobalt and urea or cyclohexanecarboxamide.
CC -!- SIMILARITY: Belongs to the nitrile hydratase subunit beta family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
```

DR EMBL; X64360; CAA45711.1; -.
DR FIR; S19715; S19715.
DR HSP; P13449; 2AHJ.
DR InterPro; IPR008990; E.transp.acc.
DR InterPro; IPR003168; NHase.beta.
DR Pfam; PF02211; NHase_beta; 1.
DR PIRSF; PIRSF001427; NHase_beta; 1.
KW Lyase.
SQ SEQUENCE 226 AA; 25201 MW; A790BCC67B319654 CRC64;
Query Match 2.8%; Score 6; DB 1; Length 226;
Best Local Similarity 100.0%; Pred.No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 164 NVVNH 169
DB 211 NVVNH 216

Search completed: August 6, 2004, 16:08:45
Job time : 15 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2004, 16:05:52 ; Search time 38 Seconds
(without alignments)
1768.563 Million cell updates/sec

Title: US-10-024-955-7
Perfect score: 213
Sequence: 1 VMKFLIAAFVAVSADPI.....VRKMTKVLAPAFKRELEKN 213

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1017041 segs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : SPREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mmc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	4.2	374	6 Q9TU81	Q9TU81 ovis aries
2	8	3.8	68	15 Q85514	Q85514 feline leuc
3	8	3.8	87	16 Q8E520	Q8E520 streptococ
4	8	3.8	16	Q8E0B3	Q8E0B3 streptococ
5	8	3.8	143	16 Q834C3	Q834C3 enterococ
6	8	3.8	318	16 Q7VTA3	Q7VTA3 bordetella
7	8	3.8	345	16 Q7WE94	Q7WE94 bordetella
8	8	3.8	345	16 Q7WZC6	Q7WZC6 bordetella
9	8	3.8	349	10 Q9SU66	Q9SU66 arabidopsis
10	8	3.8	437	2 Q9RAM9	Q9RAM9 methylobac
11	8	3.8	484	16 Q7VSN3	Q7VSN3 bordetella
12	8	3.8	528	2 Q59814	Q59814 streptomyc
13	8	3.8	632	2 Q93Q20	Q93Q20 nostoc punc
14	8	3.8	774	10 Q8S8T2	Q8S8T2 arabidopsis
15	8	3.8	918	2 Q24835	Q24835 acinetobact
16	8	3.8	1286	16 Q7UR70	Q7UR70 rhodospirell

90 P94906 microcystis
 91 Q9M08 arabidopsis
 92 Q8TY0 methanosarc
 93 Q8RX25 arabidopsis
 94 Q88GT2 pseudomonas
 95 Q828D6 streptomyces
 96 Q22780 arabidopsis
 97 Q9CAK4 arabidopsis
 98 Q9M09 caenorhabditis
 99 Q9SM59 arabidopsis
 100 Q84UT3 prunus mume

ALIGNMENTS

RESULT 1
 Q9TU81 PRELIMINARY; PRT; 374 AA.

DT 01-MAY-2000 (TREMELrel. 13, Created)
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Pulmonary surfactant-associated protein B (Fragment).
 OS Ovis aries (Sheep). Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=20215263; PubMed=10749754;
 RA Pletschmann S.M., Pison U.;
 RT "cDNA cloning of ovine pulmonary surfactant proteins A, B, and C:
 RT isolation of two different sequences for SP-B.";
 RL Am. J. Physiol. 278:L1765-L1778 (2000).
 DR EMBL; AF107544; AAF14195.1; -;
 DR HSP; P07388; 1DEF.
 DR GO; GO:0005764; C:lysosome; IEA.
 DR GO; GO:0007585; P:respiratory gaseous exchange; IEA.
 DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
 DR InterPro; IPR003119; SAPA.
 DR InterPro; IPR007856; SapB.
 DR InterPro; IPR008138; SapB_2.
 DR InterPro; IPR008140; SapB_sub.
 DR InterPro; IPR008373; Saposin.
 DR InterPro; IPR008139; SaposinB.
 DR InterPro; IPR008137; Surfactant_B.
 DR Pfam; PF02199; SAPA; 1.
 DR Pfam; PF0184; SapB_1; 1.
 DR Pfam; PF03489; SapB_2; 3.
 DR PRINTS; PR01797; SAFOSIN.
 DR ProDom; PD001732; SapB_sub; 1.
 DR ProDom; PD008002; Surfactant_B; 1.
 DR SMART; SM00162; SAPA; 1.
 DR SMART; SM00118; SAPB; 3.
 FT NON TER 1
 SQ SEQUENCE 374 AA; 41291 MW; 16A4D62804AD5A2F CRC64;

Query Match 4.2%; Score 9; DB 6; Length 374;
 Best Local Similarity 100.0%; Pred.No.2.2; 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0;

QY 188 AIFQDTVRK 196
 Db AIFQDTVRK 94

RESULT 2
 Q85514 PRELIMINARY; PRT; 68 AA.
 ID Q85514
 AC Q85514;

DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
 DE (H25-FeSV) recombination site DNA (Fragment).
 OS Feline sarcoma virus.
 OC Viruses; Retrovird viruses; Retroviridae; Mammalian type C retroviruses.
 OX NCBI_TaxID=11772;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86308237; PubMed=3018286;
 RA Besmer P., Lader B., George P.C., Bergold P.J., Qiu F.-H.,
 RA Zuckerman E.E., Hardy W.D.;
 RT "A new acute transforming feline retrovirus with fms homology
 RT specifies a C-terminally truncated version of the c-fms protein that
 RT is different from SM-Feline sarcoma virus v-fms protein.";
 RL J. Virol. 60:194-203 (1986).
 DR EMBL; M14290; AAA43043.1; -;
 DR NON TER 1
 FT NON TER 68
 SQ SEQUENCE 68 AA; 7141 MW; 1EF902F883F9ECBC CRC64;

Query Match 3.8%; Score 8; DB 15; Length 68;
 Best Local Similarity 100.0%; Pred.No.5.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 KEMTKVLA 203
 Db KEMTKVLA 9

RESULT 3

Q8E5Z0 PRELIMINARY; PRT; 87 AA.
 ID Q8E5Z0
 AC Q8E5Z0
 DT 01-MAR-2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE PtsH protein.
 GN PTH OR GBS0839.
 OS Streptococcus agalactiae (serotype III).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=216495;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NM316 / Serotype III;
 RX MEDLINE=2242508; PubMed=12354221;
 RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,
 RA Msadek T., Zouine M., Couve E., Lallouli L., Poyart C., Frieu-Cuet P.,
 RA Kunst F.;
 RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
 RT invasive neonatal disease.";
 RL Mol. Microbiol. 45:1499-1513 (2002).
 DR EMBL; AL766847; CAD46483.1; -;
 DR Sagalistic; Sbs0839; -;
 DR GO; GO:0003351; F:sugar porter activity; IEA.
 DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho...; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR001020; HPr_HisP_S.
 DR InterPro; IPR000032; HPr_protein.
 DR Pfam; PF00381; PTS-HPr; 1.
 DR PRINTS; PR00107; PHOSPHOCPPR.
 DR ProDom; PD002238; HPr_protein; 1.
 DR PROSITE; PS00369; PTS_HPR_HIS; 1.
 DR PROSITE; PS00589; PTS_HPR_SER; 1.
 DR Complete proteome.
 SQ SEQUENCE 87 AA; 8937 MW; 77D1E691843464AE CRC64;

Query Match 3.8%; Score 8; DB 16; Length 87;
 Best Local Similarity 100.0%; Pred.No.6.3; 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0;


```
QY 34 DDAAIAIE 41
Db 71 DDAAIAIE 78

RESULT 4
Q8E0B3 PRELIMINARY; PRT; 87 AA.
ID Q8E0B3
AC Q8E0B3
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Phosphocarrier protein HPr.
GN PSSH OR SAG0821.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547;
RA Tettelin H., Maignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margalit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
RA Rinaldo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
DR EMBL; AE014229; AM99708.1; -.
DR TIGR; SAG0821; -.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
DR GO; GO:0006910; P:transport; IEA.
DR InterPro; IPR001020; HPr_Hisp_S.
DR InterPro; IPR000032; HPr_protein.
DR InterPro; IPR02114; HPr_SerP_S.
DR Pfam; PF00381; PTS_HPR; 1.
DR PRINTS; PR00107; PHOSPHOCFPR.
DR PRODOM; PD002238; HPr_protein; 1.
DR PROSITE; PS00369; PTS_HPR_HIS; 1.
DR PROSITE; PS00589; PTS_HPR_SER; 1.
KW Complete proteome.
SQ SEQUENCE 87 AA; 8937 MW; 77D1E691843464AE CRC64;

Query Match 3.8%; Score 8; DB 16; Length 87;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 DDAAIAIE 41
Db 71 DDAAIAIE 78

RESULT 5
Q834C3 PRELIMINARY; PRT; 143 AA.
ID Q834C3
AC Q834C3
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN Efl745.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.

Query Match 3.8%; Score 8; DB 16; Length 318;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 VAFVAVSA 17
Db 51 VAFVAVSA 58

RESULT 7
Q7WR94 PRELIMINARY; PRT; 345 AA.
ID Q7WR94
AC Q7WR94
DT 01-OCT-2003 (TReMBLrel. 25, Created)
```

RC STRAIN=V583 / ATCC 700802;
RX MEDLINE=22550857; PubMed=12663927;
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
RA Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
RA Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
RA Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
RT "Role of mobile DNA in the evolution of vancomycin-resistant
RT Enterococcus faecalis.";
RL Science 299:2071-2074(2003).
DR EMBL; AE016952; AA081518.1; -.
DR TIGR; EF1745; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 143 AA; 15075 MW; 57626ACDAB3E5190 CRC64;

Query Match 3.8%; Score 8; DB 16; Length 143;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LIAAVAFV 13
Db 9 LIAAVAFV 16

RESULT 6
Q7VTA3 PRELIMINARY; PRT; 318 AA.
ID Q7VTA3
AC Q7VTA3
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative membrane protein.
GN BP3647.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I. / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Hoiden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitz E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrall B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; EX640422; CAE43904.1; -.
KW Complete proteome.
SQ SEQUENCE 318 AA; 33508 MW; 6BE7E3BEB340A46D CRC64;

Query Match 3.8%; Score 8; DB 16; Length 318;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 VAFVAVSA 17
Db 51 VAFVAVSA 58

RESULT 7
Q7WR94 PRELIMINARY; PRT; 345 AA.
ID Q7WR94
AC Q7WR94
DT 01-OCT-2003 (TReMBLrel. 25, Created)

```

DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative membrane protein.
GN BB0062.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Fellwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leather S., Moulé S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640437; CAE30564.1; -.
KW Complete proteome.
SQ SEQUENCE 345 AA; 36556 MW; 41DFCC0D03F507D CRC64;

Query Match 3.8%; Score 8; DB 16; Length 345;
Best Local Similarity 100.0%; Pred.No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 VAFVAUSA 17
DB 78 VAFVAUSA 85

RESULT 8
QY 10 VAFVAUSA 17
DB 78 VAFVAUSA 85

ID Q7W2C6 PRELIMINARY; PRT; 345 AA.
AC Q7W2C6;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative membrane protein.
DE BPP0062.
GN Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Fellwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leather S., Moulé S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640423; CAE39803.1; -.
KW Complete proteome.
SQ SEQUENCE 345 AA; 36556 MW; 41DFCC0D03F507D CRC64;

Query Match 3.8%; Score 8; DB 16; Length 345;
Best Local Similarity 100.0%; Pred.No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 VAFVAUSA 17
DB 78 VAFVAUSA 85

RESULT 9
QY 10 VAFVAUSA 17
DB 78 VAFVAUSA 85

ID Q9SU66 PRELIMINARY; PRT; 349 AA.
AC Q9SU66;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Hypothetical protein (YGH1-C3HC4 RING fusion protein).
GN T17F15.100 OR YGH1-RING.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Choisme N., Robert C., Brottier P., Wincker P., Cattolico L.,
RA Artiguenave F., Saurin W., Weissbach J., Mewes H.W., Mayer K.F.X.,
RA Lemke K., Schueller C., Quetier F., Salanoubat M.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Abe S., Sasaki T., Miyamoto K., Azama K., Cogburn L.A.;
RT "Structural analysis of YGH1 gene and its expression in yellowtail
RT (Seriola quinqueradiata) and phylogenetic relations to vertebrate and
RT invertebrate homologs.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AL049658; CAB41136.1; -.
DR EMBL; AB093346; BAC75820.1; -.
DR FIR; T06680; T06680.
DR InterPro; IPR007667; HIG 1 N.
DR InterPro; IPR001841; Znf ring.
DR Pfam; PF04588; HIG 1 N; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PSF0069; ZF_RING_2; 1.
DR KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 349 AA; 38686 MW; A080BD4C1D9DEBEE CRC64;

Query Match 3.8%; Score 8; DB 10; Length 349;
Best Local Similarity 100.0%; Pred.No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 QDFVVALS 143
DB 332 QDFVVALS 339

RESULT 10
QY 136 QDFVVALS 143
DB 332 QDFVVALS 339

ID Q9RAM9 PRELIMINARY; PRT; 437 AA.
AC Q9RAM9;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Homoserine dehydrogenase (EC 1.1.1.3) (Hdh).
GN Hdh.
OS Methylobacillus flagellatum.
OC Bacteria; Proteobacteria; Betaproteobacteria; Methylophilales;
OC Methylobacteriaceae; Methylobacillus.

```

```

OX NCBI_TaxID=405;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KT;
RA MEDLINE=20055654; PubMed=10589737;
RX Marchenko G.N., Marchenko N.D., Tsygankov Y.D., Chistoserov A.Y.;
RT "Organization of threonine biosynthesis genes from the obligate
RL methylophilic Methylobacillus flagellatus.";
CC Microbiology 145:3273-3282(1999).
CC -1- CATALYTIC ACTIVITY: L-HOMOSERINE + NAD(P)(+) = L-ASPARTATE 4-
CC SEMIALDEHYDE + NAD(P)H.
CC -1- SIMILARITY: BELONGS TO THE HOMOSERINE DEHYDROGENASE FAMILY.
DR EMBL; L7865; AAF21129.1; -.
DR GO; GO:0016597; F:amino acid binding; IEA.
DR GO; GO:0004412; F:homoserine dehydrogenase activity; IEA.
DR GO; GO:0004420; F:hydroxymethylglutaryl-CoA reductase (NADPH) . . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR002912; ACT.
DR InterPro; IPR002202; HMG-CoA red.
DR InterPro; IPR001342; Homoserine dh.
DR InterPro; IPR005106; NAD_binding_3.
DR Pfam; PF01842; ACT; 1.
DR Pfam; PF00742; Homoserine dh; 1.
DR Pfam; PF03447; NAD binding 3; 1.
DR PROSITE; PS00318; HMG COA REDUCTASE 2; 1.
DR PROSITE; PS01042; HOMOSER_DHCENASE; 1.
KW NADP; Oxidoreductase.
SQ SEQUENCE 437 AA; 46702 MW; B1308BFF465F9BB2 CRC64;

Query Match 3.8%; Score 8; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 DDAAIAIE 41
DB 411 DDAAIAIE 418

RESULT 11
Q7VSN3 PRELIMINARY; PRT; 484 AA.
AC Q7VSN3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glutamyl-tRNA(Leu)(Gln) amidotransferase subunit B.
GN GATB OR BP0371.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama 1 / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cordero-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX840412; CAB44703.1; -.
KW Transferase; Complete proteome.
SQ SEQUENCE 484 AA; 52332 MW; F4053691728A0E2C CRC64;

Query Match 3.8%; Score 8; DB 2; Length 528;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 PAFKRELE 211
DB 275 PAFKRELE 282

Query Match 3.8%; Score 8; DB 16; Length 484;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 IEARGLKQ 83
DB 415 IEARGLKQ 422

RESULT 12
Q59814 PRELIMINARY; PRT; 528 AA.
AC Q59814;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Principal sigma-like factor (RNA polymerase sigma factor).
GN HRDC.
OS Streptomyces aureofaciens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1894;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCM 3239;
RX MEDLINE=93083996; PubMed=1452038;
RA Kormanec J., Parkasovsky M., Potuckova L.;
RT "Four genes in Streptomyces aureofaciens containing a domain
RL characteristic of principal sigma factors.";
CC Gene 122:63-70(1992).
CC -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC THEN IS RELEASED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY.
DR EMBL; N90412; AAA26764.1; -.
DR PIR; JN0445; JN0445.
DR HSP; P00579; ISIG.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003899; F:DNA-directed RNA polymerase activity; IEA.
DR GO; GO:0016987; F:sigma factor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006740; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006352; P:transcription initiation; IEA.
DR InterPro; IPR005479; CPhase_L_D2.
DR InterPro; IPR009043; RNA pol sigma.
DR InterPro; IPR009042; Sigma70_r1_2.
DR InterPro; IPR007627; Sigma70_r2_3.
DR InterPro; IPR007624; Sigma70_r3.
DR InterPro; IPR007630; Sigma70_r4.
DR InterPro; IPR000943; Sigma70.
DR Pfam; PF04542; sigma70_r1_2; 1.
DR Pfam; PF04539; sigma70_r3; 1.
DR Pfam; PF04545; sigma70_r4; 1.
DR PRINTS; PR00046; SIGMA70FCT.
DR PROSITE; PS00867; CPhase2; 1.
DR PROSITE; PS00715; SIGMA70_1; 1.
DR PROSITE; PS00716; SIGMA70_2; 1.
KW DNA-binding; DNA-directed RNA polymerase; Sigma factor;
KW Transcription regulation; Transcription factor.
SQ SEQUENCE 528 AA; 57598 MW; A22B76FF5C065D4 CRC64;
```

RESULT 13
Q93Q20 PRELIMINARY; PRT; 632 AA.
AC Q93Q20;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PknC.
OS Nostoc punctiforme.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=63737;
RN [1]
SEQUENCE FROM N.A.
RA Wong F.C.Y., Meeks J.C.;
RT "A novel gene required for normal heterocyst differentiation pattern
in Nostoc punctiforme.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF288131; AAK68687.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Transferase.
XW
SQ SEQUENCE 632 AA; 69868 MW; C1F4321AB4970E87 CRC64;

Query Match 3.8%; Score 8; DB 2; Length 632;
Best Local Similarity 100.0%; Pred.No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 EINKAIDD 35
DB 545 EINKAIDD 552

RESULT 14
Q888T2 PRELIMINARY; PRT; 774 AA.
AC Q888T2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Subtilisin-like serine protease.
GN AT2G39850.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Rounsley S.D., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C.,
RA Sykes S.M., Kaul S., Mason T.M., Kerlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
PP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
PP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC003000; AAM14833.1; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR000209; Peptidase S8.
DR Pfam; PF00082; Peptidase S8; 1.
DR PRINTS; PRO0723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Protease.
SQ SEQUENCE 774 AA; 87729 MW; 7F570AA2FAF8DD2 CRC64;
Query Match 3.8%; Score 8; DB 10; Length 774;
Best Local Similarity 100.0%; Pred.No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 31 KAIDDAIA 38
DB 257 KAIDDAIA 264

RESULT 15
O24835 PRELIMINARY; PRT; 918 AA.
AC O24835;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Acinetobacter sp. ADP1.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ADP1.
RX MEDLINE=99287833; PubMed=10348863;
RT Segura A., Bunz P.V., D'Argenio D.A., Ornston L.N.;
RT "Genetic analysis of a chromosomal region containing vanA and vanB,
genes required for conversion of either ferulate or vanillate to
J. Bacteriol. 181:3494-3504(1999)."
DR EMBL; AF011339; AAC27114.1; -.
DR PIR; T02759; T02759.
DR InterPro; IPR008454; Cna_B.
DR Pfam; PF05738; CnaB; 2.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 918
SQ SEQUENCE 918 AA; 93381 MW; 2991F95E5B7DB94E CRC64;

Query Match 3.8%; Score 8; DB 2; Length 918;
Best Local Similarity 100.0%; Pred.No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 INKAIDDA 36
DB 78 INKAIDDA 85

RESULT 16
Q7UR70 PRELIMINARY; PRT; 1286 AA.
AC Q7UR70;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Probable myosin heavy chain.
GN RB5852.
OS Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
PP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;

DR Pfam; PF00078; rvt; 1.
 DR Pfam; PF00098; zf-CCHC; 1.
 DR SMART; SMO0343; Znf C2HC; 1.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
 DR PROSITE; PS0158; ZF_CCHC_1.
 DR Aspartyl protease; Core protein; Hydrolase; Polyprotein; Protease
 KW RNA-directed DNA polymerase; Signal; Transferase.
 FT SIGNAL 1766 1784
 FT CHAIN 75 576 POTENTIAL.
 FT CHAIN 75 576 POTENTIAL.
 FT CHAIN 202 271 POTENTIAL.
 FT CHAIN 272 519 POTENTIAL.
 FT CHAIN 520 576 POTENTIAL.
 FT CHAIN 577 1784 POTENTIAL.
 FT CHAIN 577 701 POTENTIAL.
 FT CHAIN 702 1368 POTENTIAL.
 FT CHAIN 1369 1784 POTENTIAL.
 SQ SEQUENCE 1784 AA; 200157 MW; D8162E567D054688 CRC64;

 Query Match 3.8%; Score 8; DB 15; Length 1784;
 Best Local Similarity 100.0%; Pred.No.93;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 196 KEMTKVLA 203
 |||||
 DB 513 KEMTKVLA 520

 RESULT 18
 OS089811 PRELIMINARY; PRT; 1786 AA.
 ID AC 089811;
 AC AC 089811;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 OS Gag-pol polyprotein gpr80.
 DE Feline leukemia virus.
 OC Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.
 CC NCBI_TaxID=11768;
 RN [1]_SEQUENCE FROM N.A.
 RP STRAIN=Rickard subgroup A;
 RC STRAIN=Rickard subgroup A;
 RA Chen H., Roy-Burman P.;
 RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
 CC KNOWN AS THE RETROPEPSIN FAMILY.
 DR EMBL; AF052723; AAC31801.1; -.
 DR HSSP; P03355; IMWL.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
 DR GO; GO:0019088; P:viral assembly; IEA.
 DR InterPro; IPR001969; Asprotease AS.
 DR InterPro; IPR000840; Gag MA.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Unayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtapple E., Kholdy H., Wolf A.M., Uterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Anguoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Lofthus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae";
RL Science 293:498-506(2001)
DR EMBL; AE007418; AAK75286.1; -;
DR PIR; E95136; E95136.
DR TIGR; SP1177; -;
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001020; Hpr_Hisp_S.
DR InterPro; IPR000032; Hpr_protein.
DR InterPro; IPR002114; Hpr_Serp_S.
DR Pfam; PF00381; PTS-Hpr; 1.
DR PRINTS; PR00107; PHOSPHOCHPR.
DR PRODOM; PD002238; Hpr_protein; 1.
DR TIGRFAMs; TIGR01003; PTS_Hpr_family; 1.
DR PROSITE; PS00369; PTS_HPR_HIS; 1.
DR PROSITE; PS00589; PTS_HPR_SER; 1.
KW Complete proteome.
SQ SEQUENCE 87 AA; 8939 MW; 2D610EAEF25AF70 CRC64;

Query Match 3.3%; Score 7; DB 16; Length 87;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 DDAAAI 40
Db 71 DDAAAI 77
|||||
RESULT 22
QSDPQO PRELIMINARY; PRT; 87 AA.
AC QSDPQO
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Histidine-containing phosphocarrier protein of the PTS.
DE PSSH OR SPRI1063.
GN Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OS Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAnis S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,
RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";

RL J. Bacteriol. 183:5709-5717(2001).
DR EMBL; AE008479; AAK99867.1; -;
DR PIR; E95136; E95136.
DR TIGR; SP1177; -;
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001020; Hpr_Hisp_S.
DR InterPro; IPR000032; Hpr_protein.
DR InterPro; IPR002114; Hpr_Serp_S.
DR Pfam; PF00381; PTS-Hpr; 1.
DR PRINTS; PR00107; PHOSPHOCHPR.
DR PRODOM; PD002238; Hpr_protein; 1.
DR PROSITE; PS00369; PTS_HPR_HIS; 1.
DR PROSITE; PS00589; PTS_HPR_SER; 1.
KW Complete proteome.
SQ SEQUENCE 87 AA; 8939 MW; 2D610EAEF25AF70 CRC64;

Query Match 3.3%; Score 7; DB 16; Length 87;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 DDAAAI 40
Db 71 DDAAAI 77
|||||
RESULT 23
Q99Z65 PRELIMINARY; PRT; 87 AA.
ID Q99Z65
AC Q99Z65
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE Putative phosphotransferase system phosphotransferase-containing protein
DE (His containing phosphocarrier protein).
GN PSSH OR SPY1373 OR SPY181385 OR SPY131047 OR SPY0814.
OS Streptococcus pyogenes.
OS Streptococcus pyogenes (serotype M18), and
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314, 186103, 198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=2192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.,
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX SPECIES=Streptococcus; STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=21133808; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbican K.D., Lai B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;

Query Match	3.3%;	Score 7;	DB 16;	Length 132;
Best Local Similarity	100.0%;	Pred. No. 98;		
Matches	7;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;

RESULT 26	
Q8TX30	PRELIMINARY; PRT: 142 AA.
Q8TX30	
DC	O8TX30;
AT	01-JUN-2002 (TrEMBLrel. 21, Created)
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	Membrane protein implicated in regulation of membrane protease activity.
DE	GN MK0848.
GN	OS Methanopyrus kandleri.
OS	OS Archaea; Euryarchaeota; Methanopyzi; Methanopyrales; Methanopyraceae;
OC	Methanopyrus
OC	Methanopyrus
OX	NCBI TaxID=2320;
OX	[1]
RN	SEQUENCE FROM N.A.
RP	STRAIN=AV19 / DSM 6324 / JCM 9639;
RC	MEDLINE=21927647; PubMed=11930014;
RX	Slesarev A.I., Mezheva K.V., Makarov K.S., Polushin N.N.,
RA	Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA	Natala D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA	Malykh A.G., Koonin E.V., Kozhavkin S.A.;
RA	"The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT	and monophyly of archaeal methanogens.;"
RT	Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002) .
RL	EMBL; AE010374; AAM02061.1; "
RL	GO; GO:0008233; F:peptidase activity; IEA.
DR	InterPro; IPR002810; DUF107.
DR	

DR Pfam; PF01957; NfeD; 1.
 KW Protease; Complete proteome.
 SQ SEQUENCE 142 AA; 15553 MW; 34AE784117796477 CRC64;

Query Match 3.3%; Score 7; DB 17; Length 142;
 Best Local Similarity 100.0%; Pred. No. 1.e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 KGEIGIV 99
 |||||
 DB 86 KGEIGIV 92

RESULT 27
 ID Q8KK11 PRELIMINARY; PRT; 149 AA.
 AC Q8KK11
 DT 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS ORF175.
 GN Proteus vulgaris.
 OG Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Proteus.
 OX NCBI_TaxID=585;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UR-75;
 RA Murata T., Hayashi T.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UR-75;
 RA Murata T., Hayashi T.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UR-75;
 RA Murata T., Ohnishi M., Ara T., Kaneko J., Han C.-G., Li Y.F.,
 Takashima K., Nojima H., Nakayama K., Kaji A., Kamio Y., Miki T.,
 Mori H., Ohtsubo E., Terawaki Y., Hayashi T.;
 RL "Complete Nucleotide Sequence of Plasmid Rts1: Implications for
 Evolution of Large Plasmid Genomes";
 J. Bacteriol. 184:3194-3202(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UR-75;
 RA Tian Q.B., Ohnishi M., Tabuchi A., Terawaki Y.;
 RL "A new plasmid-encoded proteic killer gene system: cloning,
 sequencing, and analyzing his locus of plasmid Rts1";
 Biochem. Biophys. Res. Commun. 220:280-284(1996).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UR-75;
 RA Janosi L., Yonemitsu H., Hong H., Kaji A.;
 RL "Molecular cloning and expression of a novel hydroxymethylcytosine-
 specific restriction enzyme (PvuRtsII) modulated by glucosylation of
 DNA";
 J. Mol. Biol. 242:45-61(1994).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UR-75;
 RA Mochida S., Tsuchiya H., Mori K., Kaji A.;
 RL "Three short fragments of Rts1 DNA are responsible for the
 temperature-sensitive growth phenotype (tsG) of host bacteria";
 J. Bacteriol. 173:2600-2607(1991).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UR-75;
 RA Nozue H., Tsuchiya K., Kamio Y.;
 RL "Nucleotide sequence and copy control function of the extension of the

RT inci region (inci-b) of Rts1.";
 RL Plasmid 19:46-56(1988).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UR-75;
 RA Tanaka M., Okawa N., Mori K., Suyama Y., Kaji A.;
 RL "Nucleotide sequence of an Rts1 fragment causing temperature-dependent
 instability";
 J. Bacteriol. 170:1175-1182(1988).
 RN [8]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UR-75;
 RA MEDLINE=85234397; PubMed=2989253;
 RA Mollet B., Clerget M., Meyer J., Iida S.;
 RL "Organization of the Tn6-related kanamycin resistance transposon
 Tn280 carrying two copies of IS26 and an IS903 variant, IS903. B.";
 J. Bacteriol. 163:55-60(1985).
 RN [9]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UR-75;
 RA Kamio Y., Tabuchi A., Itoh Y., Katagiri H., Terawaki Y.;
 RL "Complete nucleotide sequence of mini-Rts1 and its copy mutant.";
 J. Bacteriol. 158:307-312(1984).
 RN [10]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UR-75;
 RA Kamio Y., Terawaki Y.;
 RL "Nucleotide sequence of an incompatibility region of mini-Rts1 that
 contains five direct repeats";
 J. Bacteriol. 155:1185-1191(1983).
 RN [11]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UR-75;
 RA MEDLINE=68393387; PubMed=4876466;
 RA Terawaki Y., Kakizawa Y., Takayasu H., Yoshikawa M.;
 RL "Temperature sensitivity of cell growth in Escherichia coli associated
 with the temperature sensitive R(KM) factor";
 Nature 219:284-285(1968).
 DR EMBL; AP004237; BAB93737.1;
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
 DR GO; GO:0005507; F:copper ion binding; IEA.
 DR GO; GO:0005489; F:electron transporter activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR000923; BlueCu 1.
 DR PROSITE; PS00196; COPPER BLUE; 1.
 KW Hypothetical protein; Plasmid.
 SQ SEQUENCE 149 AA; 16381 MW; BDB29EB8946DE563 CRC64;

Query Match 3.3%; Score 7; DB 2; Length 149;
 Best Local Similarity 100.0%; Pred. No. 1.e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 TKVLAPA 205
 |||||
 DB 112 TKVLAPA 118

RESULT 28
 ID Q9HPE8 PRELIMINARY; PRT; 151 AA.
 AC Q9HPE8;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Vngl670c.
 GN VNG1670C.
 OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=64091;

```

[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA NG W.V., Kennedy S.P., Mahaitas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shrogha J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,
RA Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Madsen D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005076; AAG19921.1; -.
DR PIR; E84319; E84319.
KW Complete proteome.
SQ SEQUENCE 151 AA; 15766 MW; 87585DA4F36AC5AA CRC64;

Query Match 3.3%; Score 7; DB 17; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 FVVALSL 144
DB 24 FVVALSL 30

RESULT 29
Q8TLH7 PRELIMINARY; PRT; 157 AA.
AC Q8TLH7;
DT 01-JUN-2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DE Hypothetical protein MA3059.
GN MA3059.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C2A; ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932338;
RA Galagan J.E., Nuebaum C., Roy A., Endrizzi M.G., Macdonald P., Brown A.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Navior J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talanas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettnner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,
RA Pritchett M., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birken B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity."
RL Genome Res. 12:532-542(2002).
DR EMBL; AF011008; AAM06432.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 157 AA; 17167 MW; 32C8A097FA6C1602 CRC64;

Query Match 3.3%; Score 7; DB 17; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 IAAVAV 13
DB 17 IAAVAV 23

RESULT 30
Q8KGM1 PRELIMINARY; PRT; 159 AA.
AC Q8KGM1;
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DE Hypothetical conserved protein.
GN MSI288.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=57A;
RX MEDLINE=21999272; PubMed=12003951;
RA Sullivan J.T., Tzebiatowski J.R., Cruickshank R.W., Gouzy J.,
RA Brown S.D., Elliot R.M., Fleetwood D.J., McCallum N.G., Rossbach U.,
RA Stuart G.S., Weaver J.E., Webby R.J., de Bruijn F.J., Ronson C.W.;
RT "Comparative sequence analysis of the symbiosis island of
RT Mesorhizobium loti strain R7A."
RL J. Bacteriol. 184:3086-3095(2002).
DR EMBL; AL672114; CAD31320.1; -.
DR InterPro; IPR004952; DUF269.
DR Pfam; PF03270; DUF269; 1.
DR ProDom; PD008304; DUF269; 1.
KW Hypothetical protein.
SQ SEQUENCE 159 AA; 17367 MW; 873CAB6A19079519 CRC64;

Query Match 3.3%; Score 7; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 DDAAIAI 40
DB 144 DDAAIAI 150

RESULT 31
Q98AF1 PRELIMINARY; PRT; 159 AA.
AC Q98AF1;
DT 01-OCT-2001 (TREMELrel. 18, Created)
DT 01-OCT-2001 (TREMELrel. 18, Last sequence update)
DE Hypothetical protein mlr5912.
GN MLR5912.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MAF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL; AF003007; BAB52281.1; -.
DR InterPro; IPR004952; DUF269.
DR Pfam; PF03270; DUF269; 1.
DR ProDom; PD008304; DUF269; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 159 AA; 17409 MW; 3A1B615047570878 CRC64;

Query Match 3.3%; Score 7; DB 16; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 34 DDAIAAI 40
DB 144 DDAIAAI 150
|||||
7 LIAVAF 13

RESULT 32
Q878J8 PRELIMINARY; PRT; 162 AA.
AC Q878J8;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DE Conserved hypothetical protein.
GN SP51238.
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
RA Hayaishi H., Hamada S.;
RT "The genome of invasive Streptococcus pyogenes; a comparative analysis
of S. pyogenes SSI-1, SF370 and MGAS8232.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005144; BAC64333.1;
KW Hypothetical protein.
SQ SEQUENCE 162 AA; 18603 MW; 5E976308278EFCOD CRC64;

Query Match 3.3%; Score 7; DB 16; Length 162;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 FGVLSDV 185
DB 107 FGVLSDV 113
|||||

RESULT 33
ID Q89J45 PRELIMINARY; PRT; 168 AA.
AC Q89J45;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE B1r5439 protein.
GN B1r5439.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobiaceae.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriuguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shampo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005954; BAC50704.1;
KW Complete proteome.
SQ SEQUENCE 168 AA; 17805 MW; C8F979F8798C999 CRC64;

Query Match 3.3%; Score 7; DB 16; Length 168;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LIAVAF 12

QY 34 DDAIAAI 40
DB 144 DDAIAAI 150
|||||
7 LIAVAF 13

RESULT 34
Q47896 PRELIMINARY; PRT; 169 AA.
ID Q47896;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-2001 (TREMBlrel. 16, Last annotation update)
DE ORF 169.
OS Tolypothrix sp. PCC 7601.
OC Bacteria; Cyanobacteria; Nostocales; Microchaetaceae; Tolypothrix.
OX NCBI_TaxID=1188;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88260883; PubMed=2838727;
RA Parsot C., Mazel D.;
RT "Cloning and nucleotide sequence of the thrB gene from the
cyanobacterium Calothrix PCC 7601.";
RL Mol. Microbiol. 1:45-52(1987).
DR EMBL; Y00822; CA68577.1;
SQ SEQUENCE 169 AA; 19182 MW; 1A036CFA6CEA4A00 CRC64;

Query Match 3.3%; Score 7; DB 2; Length 169;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 IDDAIAA 39
DB 62 IDDAIAA 68
|||||

RESULT 35
Q8RM78 PRELIMINARY; PRT; 176 AA.
ID Q8RM78;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
OS Bacteroides fragilis.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=817;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LV25; TRANSPOSON=clv25;
RX MEDLINE=21886381; PubMed=11883096;
RA Bass K.A., Hecht D.W.;
RT "Isolation and Characterization of clv25, a Bacteroides fragilis
Chromosomal Transfer Factor Resembling Multiple Bacteroides sp.
Mobilizable Transposons.";
RL J. Bacteriol. 184:1895-1904(2002).
DR EMBL; AY053505; AAL29897.1;
KW Hypothetical protein.
SQ SEQUENCE 176 AA; 20541 MW; C608B6C2009FDE10 CRC64;

Query Match 3.3%; Score 7; DB 2; Length 176;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 AFKZELE 211
DB 13 AFKZELE 19
|||||

RESULT 36
Q9A078 PRELIMINARY; PRT; 178 AA.
ID Q9A078;
AC Q9A078;
DT 01-JUN-2001 (TREMBlrel. 17, Created)

01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
Hypothetical protein SPY0899 (Hypothetical protein spyM18_0958)
(Hypothetical protein SPYM3_0615).
SPY0899 OR SPYM18_0958 OR SPYM3_0615.
Streptococcus pyogenes, (serotype M18).
Streptococcus pyogenes (serotype M18), and
Streptococcus pyogenes (serotype M3).
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
NCBI_TaxID=1314, 186103, 198466;
[1]
SEQUENCE FROM N.A.
STRAIN=SF370 / ATCC 700294 / Serotype M1;
MEDLINE=1192684; PubMed=11926296;
Ferratti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
Perriaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
Qian Y., Jia H.G., Nayyar P.Z., Ren Q., Zhu H., Song L., White J.,
Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
"Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
[2]
SEQUENCE FROM N.A.
STRAIN=MGAS8232 / Serotype M18;
MEDLINE=21927593; PubMed=11917108;
Smoot J.C., Barbian K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
Sylvia G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
"Genome sequence and comparative microarray analysis of serotype M18
group A Streptococcus strains associated with acute rheumatic fever
outbreaks.";
Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
[3]
SEQUENCE FROM N.A.
SPECIES=S.pyogenes; STRAIN=MGAS315 / Serotype M3;
MEDLINE=21133808; PubMed=12122206;
Beres S.B., Sylvia G.L., Barbian K.D., Lei B., Hoff J.S.,
Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
Schlivert P.M., Musser J.M.;
"Genome sequence of a serotype M3 strain of group A Streptococcus:
phage-encoded toxins, the high-virulence phenotype, and clone
emergence.";
Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RL EMBL; AE006539; AA033816.1; -
DR EMBL; AE010024; AAL97599.1; -
DR EMBL; AB014149; AM79222.1; -
DR Hypothetical protein; Complete proteome.
KW
SEQUENCE 178 AA; 20463 MW; 86D29CEAC4B9703A CRC64;

Query Match 3.3%; Score 7; DB 16; Length 178;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 FGVLSDV 185
|||
DB 123 FGVLSDV 129

RESULT 37
Q8LDU7 PRELIMINARY; PRT; 183 AA.
ID Q8LDU7
AC Q8LDU7;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE Hypothetical protein.
DE Arabidopsis thaliana (Mouse-ear cress).
OS Arabidopsis
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI TaxID=3702;

RESULT 39

Q84MCS PRELIMINARY; PRT; 183 AA.
AC Q84MCS;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE A5936710.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Dale J.M., Hayashizaki Y., Heuan V.W., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones";
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT006402; AAP21210.1; -.
SQ SEQUENCE 183 AA; 20361 MW; 2B058546F077BEE3 CRC64;

Query Match 3.3%; Score 7; DB 10; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 185 VLTAIFQ 191

Db 85 VLTAIFQ 91

RESULT 40

Q9LF03 PRELIMINARY; PRT; 183 AA.
AC Q9LF03;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
GN T21H19.170.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,
RA Tabata S., Mewes H.W., Rudd S., Lemcke K., Mayer K.P.X.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL391148; CAC01866.1; -.
DR PIR; T51495; T51495.
KW Hypothetical protein.
SQ SEQUENCE 183 AA; 19958 MW; 88367943EC353AD4 CRC64;

Query Match 3.3%; Score 7; DB 10; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 185 VLTAIFQ 191

Db 87 VLTAIFQ 93

RESULT 41

Q9M878 PRELIMINARY; PRT; 185 AA.
AC Q9M878;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE F16B3.27 protein (Hypothetical protein).
GN F16B3.27 OR A73G02640.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC F16B3 genomic sequence";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation";
RL Genome Biol. 0:0-0(2002).
RN [4]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC021640; AAF32471.1; -.
DR EMBL; AY052624; AAL32702.1; -.
DR EMBL; AY088259; AAM65799.1; -.
DR EMBL; AY128811; AAMS1211.1; -.
KW Hypothetical protein.
SQ SEQUENCE 185 AA; 20382 MW; F3020E240D8C02E6 CRC64;

Query Match 3.3%; Score 7; DB 10; Length 185;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 185 VLTAIFQ 191

Db 88 VLTAIFQ 94

RESULT 42

Q8U8T1 PRELIMINARY; PRT; 190 AA.
ID Q8U8T1

```
AC Q8UBT1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Ornithine cyclodeaminase (N-terminal).
GN ARCB* OR ARU4008 OR AGR_L1691.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saepulhumachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
C58.",
RL Science 294:2317-2323 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorlo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughy D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
Agrobacterium tumefaciens C58.",
RL Science 294:2323-2328 (2001).
DR EMBL; AE009331; AAL44809.1; -;
DR EMBL; AE008283; AAK89418.1; -;
DR FIR; AC3049; AC3049.
DR PIR; H98236; H98236.
DR InterPro; IPR003462; ODC_Mu_crystall.
DR Pfam; PF02423; ODC_Mu_crystall; 1.
KW Complete proteome.
SQ SEQUENCE 190 AA; 20766 MW; AID1D6A86D0AF0FB CRC64;

Query Match 3.3%; Score 7; DB 16; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 FGVLSDV 185
DB 96 FGVLSDV 102
|||||

RESULT 43
Q894K7 PRELIMINARY; PRT; 190 AA.
AC Q894K7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Hypothetical protein.
GN CTC01534.
OS Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Massachusetts / B88;
MEDLINE=22457253; PubMed=12552129;

Query Match 3.3%; Score 7; DB 16; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 FGVLSDV 185
DB 96 FGVLSDV 102
|||||

RESULT 43
Q894K7 PRELIMINARY; PRT; 190 AA.
AC Q894K7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Hypothetical protein.
GN CTC01534.
OS Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Massachusetts / B88;
MEDLINE=22457253; PubMed=12552129;
```

```
RA Brueggemann H., Baeumer S., Fricke W.F., Wierze A., Liesegang H.,
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA Gottschalk G.;
RT "The genome sequence of Clostridium tetani, the causative agent of
tetanus disease.",
RL Proc. Natl. Acad. Sci. U.S.A. 100:13116-1321 (2003).
DR EMBL; AB015941; AAC36085.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 190 AA; 20657 MW; 403B97BDA400E05 CRC64;

Query Match 3.3%; Score 7; DB 16; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSADP 19
DB 148 VAVSADP 154
|||||

RESULT 44
Q9GUNS PRELIMINARY; PRT; 192 AA.
AC Q9GUNS;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN Y67D8C.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Bristol N2;
MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.",
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Bristol N2;
RA Edwards J., Du H., Lamar B., Kemp K., Wohldmann P., Walker C.;
RT "The sequence of C. elegans cosmid Y67D8C.",
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.",
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC025724; AAG23373.1; -;
DR WormPep; Y67D8C.7; CE22777.
KW Hypothetical protein.
SQ SEQUENCE 192 AA; 20416 MW; 05FB571237A0A9D3 CRC64;

Query Match 3.3%; Score 7; DB 5; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 FLIIAAV 10
DB 10 FLIIAAV 16
|||||

RESULT 45
Q8BE10 PRELIMINARY; PRT; 200 AA.
AC Q8BE10;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```

```
DE RHOGDI-1 (Fragment).
GN ARHGDI1 OR 5330430M07RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002)."
DR EMBL; AK075656; BAC35881.1; -.
DR GO; MGI:2178103; Arhgdia.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0005994; F:Rho GTP-dissociation inhibitor activity; IEA.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR000406; Rho GDI.
DR Pfam; PF02115; Rho GDI; 1.
DR PRINTS; PR00492; RHOGDI.
FT NON TER 1
SQ SEQUENCE 200 AA; 22948 MW; 644B94FD06AE5F8F CRC64;

Query Match 3.3%; Score 7; DB 11; Length 200;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSADP 19
DB 55 VAVSADP 61
|||||

RESULT 46
Q97VF9 PRELIMINARY; PRT; 202 AA.
ID Q97VF9
AC Q97VF9
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein SSO2667.
GN SSO2667.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweez M.J., Chan-Weiler C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erasus G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Koza C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sengen C.W., van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840 (2001).
DR EMBL; AE006861; AAK42785.1; -.
DR FPR; B90441; B90441.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 202 AA; 23604 MW; D5298B190984AE5A CRC64;

Query Match 3.3%; Score 7; DB 17; Length 202;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 116 YDLAYKL 122
DB 141 YDLAYKL 147
|||||

DE RHOGDI-1 (Fragment).
GN ARHGDI1 OR 5330430M07RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002)."
DR EMBL; AK075656; BAC35881.1; -.
DR GO; MGI:2178103; Arhgdia.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0005994; F:Rho GTP-dissociation inhibitor activity; IEA.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR000406; Rho GDI.
DR Pfam; PF02115; Rho GDI; 1.
DR PRINTS; PR00492; RHOGDI.
FT NON TER 1
SQ SEQUENCE 200 AA; 22948 MW; 644B94FD06AE5F8F CRC64;

Query Match 3.3%; Score 7; DB 11; Length 200;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VAVSADP 19
DB 55 VAVSADP 61
|||||

RESULT 46
Q97VF9 PRELIMINARY; PRT; 202 AA.
ID Q97VF9
AC Q97VF9
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein SSO2667.
GN SSO2667.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweez M.J., Chan-Weiler C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erasus G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Koza C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sengen C.W., van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840 (2001).
DR EMBL; AE006861; AAK42785.1; -.
DR FPR; B90441; B90441.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 202 AA; 23604 MW; D5298B190984AE5A CRC64;

Query Match 3.3%; Score 7; DB 17; Length 202;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 116 YDLAYKL 122
DB 141 YDLAYKL 147
|||||
```

```
RESULT 47
Q8QNN3 PRELIMINARY; PRT; 204 AA.
ID Q8QNN3
AC Q8QNN3
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Esv-1-28.
OS Ectocarpus siliculosus virus.
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phaeovirus.
OX NCBI_TaxID=37665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Esv-1;
RA Delarocue N., Bothe G., Pohl T., Knippers R., Mueller D.G., Boland W.;
RT "The complete nucleotide sequence of the Ectocarpus siliculosus virus
RT genome."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF204951; AAK1454.1; -.
DR InterPro; IPR001214; SET.
DR SMART; SM00317; SET; 1.
SQ SEQUENCE 204 AA; 22852 MW; 03D757C3A8015EE1 CRC64;

Query Match 3.3%; Score 7; DB 12; Length 204;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 VHDDIVS 113
DB 52 VHDDIVS 58
|||||

RESULT 48
Q7V089 PRELIMINARY; PRT; 212 AA.
ID Q7V089
AC Q7V089
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN PM1381.
OS Prochlorococcus marinus subsp. pastoris (strain CCMP 1378 / MED4).
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=59919;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22825698; PubMed=12917642;
RX Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinser E.R., Chisholm S.W.;
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT niche differentiation."
RL Nature 424:1042-1047 (2003).
DR EMBL; BX572093; CAE19840.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 212 AA; 24664 MW; C93EFC1175C6D136 CRC64;

Query Match 3.3%; Score 7; DB 16; Length 212;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 ITEEINK 31
DB 197 ITEEINK 203
|||||

RESULT 49
Q9UJ10
```

ID Q9UJ10 PRELIMINARY; PRT; 214 AA.
 AC Q9UJ10;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DE Photosystem II CP47 protein (Fragment).
 DE C380A1.1 (Novel protein) (Fragment).
 GN C380A1.1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wallis J.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; 257653; CAB56187.1; -.
 FT NON TER 1 214
 FT NON TER 214 214
 SQ SEQUENCE 214 AA; 23405 MW; 269F5ECC71887B96 CRC64;

Query Match 3.3%; Score 7; DB 4; Length 214;
 Best Local Similarity 100.0%; Pred.No.1.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 DDIAAI 40
 DB 146 DDIAAI 152

RESULT 50
 Q9LDP3 PRELIMINARY; PRT; 218 AA.
 AC Q9LDP3;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DE Gb|AA32471.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eucotsids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI."
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI."
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB025605; BAA98058.1; -.
 DR EMBL; AP02029; BAA97533.1; -.
 SQ SEQUENCE 218 AA; 24485 MW; 71C7AEAI09BDA44 CRC64;

Query Match 3.3%; Score 7; DB 10; Length 218;
 Best Local Similarity 100.0%; Pred.No.1.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 185 VLTAFQ 191
 DB 120 VLTAFQ 126

RESULT 51
 Q9SD24 PRELIMINARY; PRT; 225 AA.
 ID Q9SD24
 AC Q9SD24;

DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DE Photosystem II CP47 protein (Fragment).
 OS Taxus chinensis var. mairei.
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.
 OX NCBI_TaxID=120273;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lung J., Chang S.-H., Tsai C.-J., Ho C.-K.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY043261; AAK85713.1; -.
 DR GO; GO:0009507; C:chloroplast; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0009521; C:photosystem; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR000932; PSIIProt.
 DR Pfam; PF00421; PSII; 1.
 KW Chloroplast.
 FT NON TER 1 1
 FT NON TER 225 AA; 25060 MW; CD1CDC551E6789E3 CRC64;
 SQ SEQUENCE 225 AA; 25060 MW; CD1CDC551E6789E3 CRC64;

Query Match 3.3%; Score 7; DB 8; Length 225;
 Best Local Similarity 100.0%; Pred.No.1.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 EGVKKA 101
 DB 90 EGVKKA 96

RESULT 52
 Q9HE10 PRELIMINARY; PRT; 250 AA.
 ID Q9HE10
 AC Q9HE10;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Putative guanine nucleotide binding protein similar to yor223w.
 GN SPAC20H4.02.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972h-;
 RA Seeger K., Harris D., Wood V., Rajandream M.A., Barrell B.G.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL512487; CAC19732.1; -.
 DR GenDB SPombe; SPAC20H4.02; -.
 SQ SEQUENCE 250 AA; 28262 MW; 1AD18FD0D28F3E63 CRC64;

Query Match 3.3%; Score 7; DB 3; Length 250;
 Best Local Similarity 100.0%; Pred.No.1.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 LSLEISD 148
 DB 22 LSLEISD 28

RESULT 53
 Q9VC67 PRELIMINARY; PRT; 252 AA.
 ID Q9VC67
 AC Q9VC67; Q8SZB4;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE CG13618-PA (R08075p).

GN CGL13618.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 CX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Anantides P.G., Scher S.E., Richards S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richardson M., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.H., Blazeg R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abriel J.P., Agbayani A., An H.D., Andrews-Pfannkuch C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Borchan M.R., Buck J., Brokstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Posler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lai X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodagef, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2195-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Anantides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Balgwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kanink B., Prochnik S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;

RT "Annotation of Drosophila melanogaster genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RX Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Krommiller B., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003748; AAF56308.2; -;
 DR EMBL; AY070987; AAL48609.1; -;
 DR FlyBase; FBgn0039203; CGL13618.
 DR InterPro; IPR004272; Odorant_binding.
 DR Pfam; PF03027; DUF233; 1.
 DR SMART; SM00700; JHBP; 1.
 SQ SEQUENCE 252 AA; 28845 MW; E7442C3888E65A5A CRC64;
 Query Match 3.3%; Score 7; DB 5; Length 252;
 Best Local Similarity 100.0%; Pred.No. 1.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 LLIAAVA 11
 DB 11 LLIAAVA 17
 RESULT 54
 Q8CV69 PRELIMINARY; PRT; 258.AA.
 AC Q8CV69;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical conserved protein.
 GN O80888
 OS Oceanobacillus iheyensis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
 OX NCBI_TaxID=182710;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HT831 / DSM 14371 / JCM 11309;
 RX MEDLINE=2220767; PubMed=12235376;
 RA Takami H., Takaki Y., Uchiyama I.;
 RT "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya
 RT Ridge and its unexpected adaptive capabilities to extreme
 RT environments."
 RL Nucleic Acids Res. 30:3927-3935(2002).
 DR EMBL; AP004596; BAC12844.1; -;
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002198; ADH_short.
 DR PROSITE; PS00061; ADH_SHORT; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 258 AA; 29547 MW; 0A29027062C7163C CRC64;
 Query Match 3.3%; Score 7; DB 16; Length 258;
 Best Local Similarity 100.0%; Pred.No. 1.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 172 LSILDPI 178
 DB 30 LSILDPI 36

```
DR GO: 0005800; P: oxygen and reactive oxygen species metabolism; IEA.
DR InterPro: IPR000345; CytC_heme_BS.
DR InterPro: IPR001989; Radical_activat.
DR InterPro: IPR007197; Radical_SAM.
DR Pfam: PF04055; Radical_SAM; 1.
DR ProDom: PD004758; Radical_activat; 1.
DR PROSITE: PS00190; CYTOCHROME_C; 1.
KW Lyase.
SQ SEQUENCE 270 AA; 29014 MW; 57F5C20CD2BE2465 CRC64;

Query Match 3.3%; Score 7; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 LAYKLG 124
DB 240 LAYKLG 246

RESULT 57
O98026 PRELIMINARY; PRT; 270 AA.
AC O98026;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE MHC class I chain-related protein (Fragment).
GN MIC.
OS Hylobates lar (Common gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=9580;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99261547; PubMed=10331162;
RA Pellet F.; Vaneensberghe C.; Debre P.; Sumyuen M.H., Theodorou I.;
RT "MIC genes in non-human primates.";
RL Eur. J. Immunogenet. 26:239-241(1999).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
MICROGLOBULIN) (BY SIMILARITY).
CC EMBL: AF045604; AD11614.1; -.
DR GO: 0016021; C: integral to membrane; IEA.
DR GO: 0006955; P: immune response; IEA.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003597; Ig_c1.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001039; MHC_I.
DR Pfam: PF00447; Ig; 1.
DR Pfam: PF00129; MHC_I; 1.
DR ProDom: PD000050; MHC_I; 1.
DR SMART: SM00407; IGc1; 1.
DR PROSITE: PS00835; IG_LIKE; 1.
DR PROSITE: PS00290; IG_MHC; 1.
DR GlycoProtet; Transmembrane.
FT NON_TER 1
FT NON_TER 270
SQ SEQUENCE 270 AA; 31026 MW; AAD5DC05F51DBESF CRC64;

Query Match 3.3%; Score 7; DB 7; Length 270;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 EGNITWT 155
DB 190 EGNITWT 196

RESULT 58
O9D9W2 PRELIMINARY; PRT; 282 AA.
ID O9D9W2
AC O9D9W2;
```

```
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 1700026N20Rik protein.
GN 1700026N20Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gisi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schram L.M., Staubli F., Suzuki R., Tonita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., King B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:695-690(2001).
DR EMBL; AK006398; BAB24568.1; -.
DR HSSP; P28867; 1PTQ.
DR MGD; MGI:191743; 1700026N20Rik.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR002219; DAG_PE-Bind.
DR InterPro; IPR000198; RhoGAP.
DR InterPro; IPR008936; Rho GAP.
DR Pfam; PF00130; DAG_PE-Bind; 1.
DR Pfam; PF00620; RhoGAP; 1.
DR PRINTS; PR00008; DAGPEDOMAIN.
DR SMART; SM00109; C1; 1.
DR SMART; SM00324; RhoGAP; 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR PROSITE; PS00081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS00238; RHO GAP; 1.
SQ SEQUENCE 282 AA; 32540 MW; C53D6D40291F4EF0 CRC64;

Query Match 3.3%; Score 7; DB 11; Length 282;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 76 IEARGLK 82
Db 115 IEARGLK 121
|||||

RESULT 59
Q97ZG6 PRELIMINARY; PRT; 286 AA.
AC Q97ZG6;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein SSO0950.
GN SSO0950.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
CX NCBI_TaxID=2287;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erasus G., Fletcher C., Gordon P.M.K., Medina N., Peng X.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Theriault C., Tolstrup N.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; AE006715; AAK41224.1; -.
DR PIR; A90246; A90246.
DR InterPro; IPR002731; ATPase_BadF.
DR Pfam; PF01869; BcrAD_BadFG; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 285 AA; 32278 MW; 71D00DBE5D978028 CRC64;

Query Match 3.3%; Score 7; DB 17; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 206 FKRELEK 212
Db 259 FKRELEK 265
|||||

RESULT 60
Q9R468 PRELIMINARY; PRT; 289 AA.
AC Q9R468;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Ornithine cyclodeaminase.
GN TORE21.
OS Agrobacterium tumefaciens.
OG Plasmid pTiC58.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=358;
RN [1]
RP SEQUENCE FROM N.A.
RA Gielen J., Terryn N., Van Montagu M., Villarroel R.;
RT "Complete nucleotide sequence of the T-DNA region of the plant tumor
RT inducing Agrobacterium tumefaciens Ti plasmid pTiC58.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ237588; CAB44645.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR InterPro; IPR003462; ODC_Mu_crystall.
DR Pfam; PF02423; ODC_Mu_crystall; 1.
KW Plasmid.
SQ SEQUENCE 289 AA; 31620 MW; 905A9B13022B9E5D CRC64;

Query Match 3.3%; Score 7; DB 2; Length 289;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 179 FGVLSDV 185
Db 31 FGVLSDV 37
|||||

RESULT 61
Q8TVG6 PRELIMINARY; PRT; 289 AA.
AC Q8TVG6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Uncharacterized protein.
GN MK1423.
```

OS Methanopyrus kandleri.
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
 OC Methanopyrus.
 OX NCBI_TaxID=2320;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AV19 / DSM 6324 / JCM 9639;
 RA MEDLINE=21927647; PubMed=11930014;
 RX Slesarev A.I., Mezheva V.V., Makarova K.S., Polushin N.N.,
 RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
 RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
 RA Malykh A.G., Koonin E.V., Kozlovskiy S.A.;
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
 RT and monophyly of archaeal methanogens.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
 DR EMBL; AE010434; AA02636.1; -;
 KW Complete proteome.
 SQ SEQUENCE 289 AA; 30485 MW; 9F34178D2E6366FB CRC64;

Query Match 3.3%; Score 7; DB 17; Length 289;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 LSDVLTA 188
 |||||
 Db 95 LSDVLTA 101

RESULT 62
 Q8A0V9 PRELIMINARY; PRT; 290 AA.
 ID Q8A0V9
 AC Q8A0V9;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Conserved hypothetical protein.
 GN BT3912.
 OS Bacteroides thetaiotaomicron.
 OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
 OC Bacteroidaceae; Bacteroides.
 OX NCBI_TaxID=818;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VPI-5482 / ATCC 29148;
 RX MEDLINE=22550858; PubMed=12663928;
 RA Xu J., Bjursell M.K., Hamrod J., Deng S., Carmichael L.K.,
 RA Chiang H.C., Hooper L.V., Gordon J.I.;
 RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
 RL Science 299:2074-2076(2003).
 DR EMBL; AE016942; AA079017.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 290 AA; 33418 MW; 55E72428D896C4B CRC64;

Query Match 3.3%; Score 7; DB 16; Length 290;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LLIAVA 11
 |||||
 Db 6 LLIAVA 12

RESULT 63
 Q8VLX7 PRELIMINARY; PRT; 291 AA.
 ID Q8VLX7
 AC Q8VLX7;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE CsrB protein.
 GN C2RB.
 OS Thermus thermophilus.
 OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;

Query Match 3.3%; Score 7; DB 2; Length 292;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OC Thermus.
 OX NCBI_TaxID=274;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KT8;
 RA Spada S., Pembroke J.T., Wall J.G.;
 RT Cloning and characterisation of the csrB metal cation efflux protein
 RT from T. thermophilus.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ307316; CAC83722.1; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008324; P:cation transporter activity; IEA.
 DR GO; GO:0006812; P:cation transport; IEA.
 DR InterPro; IPR002524; Cation_efflux.
 DR Pfam; PF01545; Cation_efflux; 1.
 DR TIGRFAMs; TIGR01297; CDF; 1.
 SQ SEQUENCE 291 AA; 31233 MW; 21CBA61D9DC8FB73 CRC64;

Query Match 3.3%; Score 7; DB 2; Length 291;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 VLSDVLT 187
 |||||
 Db 147 VLSDVLT 153

RESULT 64
 Q9W33 PRELIMINARY; PRT; 292 AA.
 ID Q9W33
 AC Q9W33;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative ABC transporter ATP-binding subunit.
 DE PENJ.
 GN OS
 OS Pedicoccus pentosaceus.
 OG Plasmid pMD136.
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae; Pedicoccus.
 OX NCBI_TaxID=1255;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC43200;
 RA Giacomini A., Marazzan G., Salvato P., Squartini A., Nuti M.P.;
 RT "Nucleotide sequence of plasmid pMD136 from Pedicoccus pentosaceus
 RT ATCC43200.";
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC43200;
 RA Kantor A., Mett A., Shapira R.;
 RT "Pedicoccus pentosaceus pediocin A encoding plasmid, pMD136.";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF033858; RAD39627.1; -;
 DR EMBL; AF069302; RAD25904.1; -;
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0003524; F:ATP binding; IEA.
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
 DR GO; GO:000166; F:nucleotide binding; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003439; ABC_transporter.
 DR Pfam; PF00005; ABC_tran; 1.
 DR ProDom; PD000006; ABC_transporter; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
 KW ATP-binding; Plasmid.
 SQ SEQUENCE 292 AA; 32454 MW; 7A7F3B3EAE2B6047 CRC64;

Query Match 3.3%; Score 7; DB 2; Length 292;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 143 SLEISDE 149
Db 23 SLEISDE 29

RESULT 65
Q8GIX8 PRELIMINARY; PRT; 296 AA.
AC Q8GIX8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mycoplasma sp. 'bovine group 7'.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2105;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG50;
RA Djordjevic S.P., Vilei E.M., Frey J.;
RT "Cloning and characterisation of a 7.9-kilobase chromosomal region of
RT Mycoplasma sp. bovine group 7 (PG50) encoding the glycerol transport
RT locus (GSAABC) and several putative membrane proteins: Implications
RT for virulence and species designation.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ419906; CAD12049.1; -.
DR InterPro; IPR005046; DUF285.
DR Pfam; PF03382; DUF285; 1.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 296 AA; 32869 MW; 304FAE8D061407ED CRC64;

Query Match 3.3%; Score 7; DB 2; Length 296;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 AFVAUSA 17
Db 270 AFVAUSA 276

RESULT 66
Q8IDW9 PRELIMINARY; PRT; 297 AA.
AC Q8IDW9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN Mal13P1.180.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Beriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrall B.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL84509; CAD52500.1; -.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 297 AA; 35497 MW; 55337BFC6DCD2DF0 CRC64;

Query Match 3.3%; Score 7; DB 5; Length 297;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 ITEINK 31
Db 21 ITEINK 27

RESULT 67
```

```
Q84ZH2 PRELIMINARY; PRT; 297 AA.
AC Q84ZH2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative class III acidic chitinase.
GN P0656C04.2.
OS Oryza sativa (japonica cultivar-group). Embryophyta; Tracheophyta;
OC Eukaryota; Viridiplantae; Streptophyta; Magnoliophyta; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
RT clone:P0656C04.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF004346; BAC55717.1; -.
DR GO; GO:0016787; P:hydrolase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
DR Pfam; PF0704; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE 18; 1.
DR PROSITE; PS01095; CHITINASE 18; 1.
SQ SEQUENCE 297 AA; 31599 MW; EA0F0F306CF19C51 CRC64;

Query Match 3.3%; Score 7; DB 10; Length 297;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LLIAAVA 11
Db 11 LLIAAVA 17

RESULT 68
Q96169 PRELIMINARY; PRT; 300 AA.
AC Q96169;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RNA-binding protein (KH domain).
GN PFB0370C.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99021743; PubMed=9804551;
RA Gardner M.J., Tettelin H., Carucci D.J., Cummings L.M., Aravind L.,
RA Koonin E.V., Shalton S., Mason T., Yu K., Fujii C., Pederson J.,
RA Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pertea M.,
RA Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Science 282:1126-1132(1998).
DR EMBL; AE001390; AAC71863.1; -.
DR PIR; F71616; F71616.
DR GO; GO:0003676; P:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH_dom.
DR SMART; SM00322; KH; 1.
SQ SEQUENCE 300 AA; 36075 MW; 079F60A811CC3242 CRC64;

Query Match 3.3%; Score 7; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 207 KRELEKN 213
```


RC STRAIN=CO-92 / Biovar Orientalis;
RA MEDLINE=21470413; PubMed=11586360;
RX Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moulis S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of *Yersinia pestis*, the causative agent of plague";
RL Nature 413:523-527(2001).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fotherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of *Yersinia pestis* KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
DR EMBL; AJ414451; CAC90808.1; -;
DR EMBL; AE013833; AAM85873.1; -;
DR PIR; AD0243; AD0243.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 310 AA; 35163 MW; D8BAC4022BF9B396 CRC64;

Query Match 3.3%; Score 7; DB 16; Length 310;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 181 VLSDVLT 187
Db 278 VLSDVLT 284
|||||

RESULT 73
Q8Y307 PRELIMINARY; PRT; 313 AA.
AC Q8Y307; PubMed=11823852;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative oxidoreductase protein (EC 1.1.1.-).
GN RSC0174 OR RS01045.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GM11000;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Sigulier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*.";
RL Nature 415:497-502(2002).
DR EMBL; AL646057; CAD13702.1; -;
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000620; DUF6.
DR InterPro; IPR000566; Lipoclin_cytFABP.
DR Pfam; PF00892; DUF6; 1.
DR PROSITE; PS00213; LIPOCALIN; 1.
KW Plasmid; Complete proteome.
SQ SEQUENCE 316 AA; 34411 MW; 08F22D71759F791D CRC64;

Query Match 3.3%; Score 7; DB 16; Length 316;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LIAAVAF 12
Db 33 LIAAVAF 39
|||||

RESULT 75
Q9N549 PRELIMINARY; PRT; 324 AA.
AC Q9N549;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Serpentine receptor, class I protein 40.
GN Y27F2A.3 OR SRI-40.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Bristol N2;
RM MEDLINE=99069613; PubMed=9851916;

Query Match 3.3%; Score 7; DB 16; Length 313;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 DDAAIAI 40
Db 83 DDAAIAI 89
|||||

RESULT 74
Q8XTH7 PRELIMINARY; PRT; 316 AA.
AC Q8XTH7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Probable transmembrane protein.
GN RSP0133 OR RS02987.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OG Plasmid megaplasmid.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GM11000;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Sigulier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*.";
RL Nature 415:497-502(2002).
DR EMBL; AL646076; CAD17284.1; -;
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000620; DUF6.
DR InterPro; IPR000566; Lipoclin_cytFABP.
DR Pfam; PF00892; DUF6; 1.
DR PROSITE; PS00213; LIPOCALIN; 1.
KW Plasmid; Complete proteome.
SQ SEQUENCE 316 AA; 34411 MW; 08F22D71759F791D CRC64;

Query Match 3.3%; Score 7; DB 16; Length 316;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LIAAVAF 12
Db 33 LIAAVAF 39
|||||

RESULT 75
Q9N549 PRELIMINARY; PRT; 324 AA.
AC Q9N549;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Serpentine receptor, class I protein 40.
GN Y27F2A.3 OR SRI-40.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Bristol N2;
RM MEDLINE=99069613; PubMed=9851916;

```
RA Wilson R.;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Cotton M., Graves T.;
RT "The sequence of C. elegans cosmid Y27F2A.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.H.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006730; AAF60478.3; -.
DR WormPep; Y27F2A.3; CB33728.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004930; F:G-protein coupled receptor activity; IEA.
DR GO; GO:0004872; P:receptor activity; IEA.
DR InterPro; IPR003003; 7TM_chemrecept.
DR InterPro; IPR000168; Nm7TM_chemrecept.
DR Pfam; PF01604; 7tm_5; 1.
SQ SEQUENCE 324 AA; 36880 MW; 0EC7940AA64CAF98 CRC64;

Query Match 3.3%; Score 7; DB 5; Length 324;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 FLLIAAV 10
Db 191 FLLIAAV 197

RESULT 76
Q98LI1 PRELIMINARY; PRT; 331 AA.
AC Q98LI1;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Permease protein of sugar ABC transporter.
GN MLL1013.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-335(2000).
DR EMBL; AF002996; BAB48482.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001851; Bac_inmem_transp.
DR Pfam; PF02653; BPD_transp_2; 1.
KW Complete Proteome.
SQ SEQUENCE 331 AA; 34057 MW; 5C83CC386FEA9E3A CRC64;

Query Match 3.3%; Score 7; DB 16; Length 331;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 IAAVAV 13
Db 188 IAAVAV 194

RESULT 77
Q80XD1 PRELIMINARY; PRT; 332 AA.
AC Q80XD1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to chimerin (Chimaerin) 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Straussberg R.;
RA Straussberg R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC051139; AAH51139.1; -.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR000198; RhoGAP.
DR InterPro; IPR008936; Rho GAP.
DR Pfam; PF00130; DAG_PE-bind; 1.
DR Pfam; PF00620; RhoGAP; 1.
DR PRINTS; PR00008; DAGPEDOMAIN.
DR SMART; SM00109; Ci; 1.
DR SMART; SM00324; RhoGAP; 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR PROSITE; PS00081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS0238; RHO GAP; 1.
SQ SEQUENCE 332 AA; 38176 MW; E35B94CCEFE3FDFB CRC64;

Query Match 3.3%; Score 7; DB 11; Length 332;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 IEARGLK 82
Db 165 IEARGLK 171

RESULT 78
Q7TOX4
```



```
ID AC Q7TOX4 PRELIMINARY; PRT; 332 AA.
AC Q7TOX4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Maman A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC055998; AAH55998.1; -.
KW Hypothetical protein.
SQ SEQUENCE 332 AA; 37724 MW; 41A4C75656C4D513 CRC64;

Query Match 3.3%; Score 7; DB 13; Length 332;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 169 IGGLSIL 175
Db 203 IGGLSIL 209

RESULT 79
ID Q9WYW8 PRELIMINARY; PRT; 333 AA.
AC Q9WYW8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Oligopeptide ABC transporter, ATP-binding protein.
GN TM0498.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
```

```
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linner K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Maman A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC055998; AAH55998.1; -.
KW Hypothetical protein.
SQ SEQUENCE 332 AA; 37724 MW; 41A4C75656C4D513 CRC64;

Query Match 3.3%; Score 7; DB 15; Length 333;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 95 EBGIVKA 101
Db 20 EBGIVKA 26

RESULT 80
Q97XA9 PRELIMINARY; PRT; 345 AA.
AC Q97XA9;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Predicted phosphohydrolase, Icc family.
GN CAC1010.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838 (2001).
DR EMBL; AF007616; AAK78986.1; -.
DR PIR; G97024; G97024.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR004843; M:ppetrase.
DR Pfam; PF00149; Metallophos; 1.
KW Hydrolase; Complete proteome.
```

```

SQ SEQUENCE 345 AA; 38556 MW; D8D4680FE6D4932B CRC64;
  Query Match 3.3%; Score 7; DB 16; Length 345;
  Best Local Similarity 100.0%; Pred. No. 2.3e+02;
  Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LLIAAVA 11
DB 10 LLIAAVA 16
|||||

RESULT 81
Q3SMS3 PRELIMINARY; PRT; 349 AA.
AC Q3SMS3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN T30210.70 OR AT4G09310.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Van Der Schueren J., Chuang Y.J., Voet M., Robben J.,
RA Volkart G., Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C.;
RA Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Robben J., Grymonprez B., Volckaert G., Spiegel L.A., Huang E.N.,
RA Nascimento L.U., de la Bastide M., Vil D.M., Preston R.R., Madero A.,
RA Shah R., O'Shaughnessy A., Rodriguez M., Shekher M., Schutz K.,
RA See L.H., Swaby I., Habermann K., Dedhia N.N., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL117386; CAB55694.1; -
DR EMBL; AL161514; CAB78054.1; -
DR PIR; T17130; T17130.
DR InterPro; IPR006594; Lish.
DR InterPro; IPR003877; SPRY_receptor.
DR Pfam; PF00622; SPRY; 1.
DR SMART; SM00667; Lish; 1.
DR PROSITE; PS00896; Lish; 1.
KW Hypothetical protein.
SQ SEQUENCE 345 AA; 38892 MW; 990E4852C9F18B5E CRC64;
  Query Match 3.3%; Score 7; DB 10; Length 349;
  Best Local Similarity 100.0%; Pred. No. 2.3e+02;
  Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 IDDAIAA 39
DB 265 IDDAIAA 271
|||||

RESULT 82
Q44332 PRELIMINARY; PRT; 351 AA.
AC Q44332;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

```

```

DE Ornithine cyclodeaminase.
GN ARCB.
OS Agrobacterium tumefaciens.
OG Plasmid pAKR10.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OC NCBI_TaxID=358;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R10;
RX MEDLINE=96178949; PubMed=8606160;
RA Cho K., Fuqua C., Martin B.S., Winans S.C.;
RT "Identification of Agrobacterium tumefaciens genes that direct the
RT complete catabolism of octopine.";
RL J. Bacteriol. 178:1872-1880(1996).
DR EMBL; U39262; AAC43978.1; -
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR InterPro; IPR003462; ODC_Mu_crystall.
DR Pfam; PF02423; ODC_Mu_crystall; 1.
KW Plasmid.
SQ SEQUENCE 351 AA; 38683 MW; 7F074C5024A4F39F CRC64;
  Query Match 3.3%; Score 7; DB 2; Length 351;
  Best Local Similarity 100.0%; Pred. No. 2.3e+02;
  Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 FGVLSDV 185
DB 96 FGVLSDV 102
|||||

RESULT 83
Q9RG93 PRELIMINARY; PRT; 354 AA.
AC Q9RG93;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE TIOF188 protein.
DE TIOF188.
OS Agrobacterium tumefaciens.
OG Plasmid pTi-SAKURA.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OC NCBI_TaxID=358;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RX MEDLINE=20184752; PubMed=10721727;
RA Suzuki K., Hattori Y., Uraji M., Ohta N., Iwata K., Murata K.,
RA Katoh A., Yoshida K.;
RT "Complete nucleotide sequence of a plant tumor-inducing Ti plasmid.";
RL Gene 242:331-336(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RX MEDLINE=98193120; PubMed=9524202;
RA Suzuki K., Ohta N., Hattori Y., Uraji M., Katoh A., Yoshida K.;
RT "Novel structural difference between nopaline- and octopine- type trbJ
RT gene: construction of genetic and physical map and sequencing of
RT trbJ/traI and rep gene clusters of a new Ti plasmid pTi-SAKURA.";
RL Biochim. Biophys. Acta 1396:1-7(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RA Hattori Y., Suzuki K., Ohta N., Uraji M., Katoh A., Yoshida K.;
RT "Genome structure of pTi-SAKURA(I): Strategy for DNA sequencing of a
RT Japanese cherry-Ti plasmid.";
RL Nucleic Acids Symp. Ser. 37:159-160(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RA Ohta N., Suzuki K., Hattori Y., Uraji M., Katoh A., Yoshida K.;

```


Best Local Similarity 100.0%; Pred. No. 2.4e+02; Mismatches 0; Indels 0; Gaps 0;

QY 35 DAIAAIE 41
 |||||
 Db 31 DAIAAIE 37

RESULT 87

Q9FOG8 PRELIMINARY; PRT; 364 AA.
 AC Q9FOG8; (TREMELrel. 16, Created)
 DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Endoglucanase precursor.
 GN ENDS.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP Michael P., Belaich A., Courtois B., Courtois J.;
 RA "Cloning sequencing and overexpression of a sinorhizobium meliloti
 RT endoglucanase gene";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF233448; AAG44364.1;
 DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . . ; IEA.
 DR GO; GO:0005595; F:carbohydrate metabolism; IEA.
 DR InterPro; IPR001547; Glyco_hydro_5.
 DR Pfam; PF00150; cellulase; 1.
 DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
 KW Signal.
 FT SIGNAL.
 FT CHAIN 1 27 POTENTIAL.
 FT CHAIN 28 364 ENDOGLUCANASE.
 SQ SEQUENCE 364 AA; 39832 MW; C927C770AAE832AF CRC64;

Query Match 3.3%; Score 7; DB 2; Length 364;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 DAIAAIE 41
 |||||
 Db 259 DAIAAIE 265

RESULT 88

Q93GC2 PRELIMINARY; PRT; 374 AA.
 AC Q93GC2; (TREMELrel. 19, Created)
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
 DE GLVA.
 GN GLVA.
 OS Xanthomonas campestris.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=339;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP MEDLINE=21417034; PubMed=11526012;
 RA Heu S., Oh J., Kang Y., Ryu S., Cho S.K., Cho Y., Cho M.;
 RT "Gly Gene cloning and expression and purification of glycinecin A, a
 bacteriocin produced by Xanthomonas campestris pv. glycines 8ra.";
 RL Appl. Environ. Microbiol. 67:4105-4110(2001).
 DR EMBL; AF281069; AAL02153.1;
 SQ SEQUENCE 374 AA; 41915 MW; 7E0B20BE98204A8A CRC64;

Query Match 3.3%; Score 7; DB 2; Length 374;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 DVLTAIF 190
 |||||
 Db 12 DVLTAIF 18

RESULT 89

Q8DIQ3 PRELIMINARY; PRT; 375 AA.
 AC Q8DIQ3; (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE Tlr1528 protein.
 GN TLR1528.
 OS Synecococcus elongatus (Thermosynechococcus elongatus).
 OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
 OX NCBI_TaxID=32046;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=BP-1;
 RX MEDLINE=2225144; PubMed=12240834;
 RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
 RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
 RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the thermophilic cyanobacterium
 Thermosynechococcus elongatus BP-1";
 RL DNA Res. 9:123-130(2002).
 DR EMBL; AF005374; BAC09080.1;
 KW Complete proteome.
 SQ SEQUENCE 375 AA; 42768 MW; 551F0668557ED14B CRC64;

Query Match 3.3%; Score 7; DB 16; Length 375;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LLIAAVA 11
 |||||
 Db 24 LLIAAVA 30

RESULT 90

P94906 PRELIMINARY; PRT; 376 AA.
 AC P94906; (TREMELrel. 03, Created)
 DT 01-MAY-1997 (TREMELrel. 03, Last sequence update)
 DT 01-MAY-1997 (TREMELrel. 03, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE RpoD2 sigma factor (RNA polymerase sigma factor).
 GN RpoD2.
 OS Microcystis aeruginosa.
 OC Bacteria; Cyanobacteria; Chroococcales; Microcystis.
 OX NCBI_TaxID=1126;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=K-81;
 RX MEDLINE=9723629; PubMed=9116041;
 RA Asayama M., Suzuki A., Nozawa S., Yamada A., Tanaka K., Takahashi H.,
 RA Aida T., Shirai M.;
 RT "A new sigma factor homolog in a cyanobacterium: cloning, sequencing,
 RT and light-responsive transcripts of rpoD2 from Microcystis aeruginosa
 K-81";
 RL Biochim. Biophys. Acta 1351:31-36(1997).
 CC -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
 ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
 THEN IS RELEASED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY.
 DR EMBL; D86575; BAAL13122.1;
 DR HSSP; P00579; 1SIG.
 DR GO; GO:0003899; F:DNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0016987; F:sigma factor activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.

```

DR GO: GO:0016740; P:transferrase activity; IEA.
DR GO: GO:0006355; P:regulation of transcription; DNA-dependent; IEA.
DR GO: GO:0006352; P:transcription initiation; IEA.
DR InterPro: IPR009043; RNA_pol_sigma.
DR InterPro: IPR009042; Sigma70_r1.2.
DR InterPro: IPR007627; Sigma70_r2.
DR InterPro: IPR007624; Sigma70_r3.
DR InterPro: IPR007630; Sigma70_r4.
DR InterPro: IPR000943; Sigma70.
DR Pfam: PF00140; sigma70_r1_2; 1.
DR Pfam: PF04542; sigma70_r2; 1.
DR Pfam: PF04539; sigma70_r3; 1.
DR Pfam: PF04545; sigma70_r4; 1.
DR PRINTS: PR00046; SIGMA70_F0CT.
DR PROSITE: PS00716; SIGMA70_2; 1.
KW DNA-binding; DNA-directed RNA polymerase; Sigma factor;
KW Transcription regulation; Transferrase.
SQ SEQUENCE 376 AA; 43784 MW; 9400034BD91E3D06 CRC64;

Query Match 3.3%; Score 7; DB 2; Length 376;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 VLSDVLT 187
DB 310 VLSDVLT 316

RESULT 91
Q9MQ08 PRELIMINARY; PRT; 377 AA.
AC Q9MQ08;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN Arabidopsis thaliana (Mouse-ear cross).
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; rosids;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
CX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Robben J., Grymonprez B., Volckaert G., Spiegel L.A., Huang E.N.,
RA Nascimento L.U., de la Bastide M., Vil D.M., Preston R.R., Matero A.,
RA Shah R., O'Shaughnessy A., Rodriguez M., Shekher M., Schutz K.,
RA See L.H., Swaby I., Habermann K., Dedhia N.N., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL161514; CAB78044.1; -.
DR InterPro: IPR006595; CTLL.
DR InterPro: IPR006594; Lish.
DR InterPro: IPR003877; SPRY_receptor.
DR Pfam: PF00622; SPRY; 1.
DR SMART: SM00668; CTLL; 1.
DR SMART: SM00667; Lish; 1.
DR PROSITE: PS0897; CTLL; 1.
DR PROSITE: PS0896; Lish; 1.
KW Hypothetical protein.
SQ SEQUENCE 377 AA; 41876 MW; 03156BC0C804DB62 CRC64;

```

```

Query Match 3.3%; Score 7; DB 10; Length 377;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 IDDAIAA 39
DB 274 IDDAIAA 280

RESULT 92
Q8TRYO PRELIMINARY; PRT; 377 AA.
AC Q8TRYO;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein MAL031.
GN MAL031.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
CX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Navlor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talanas J., Titrrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Krzycki J.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettner H.C., Ziericke J.A., Smith K.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birken B.;
RA "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE010769; AAMB4461.1; -.
DR GO: GO:0005489; P:electron transporter activity; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR InterPro: IPR001450; 4Fe4S_ferredoxin.
DR InterPro: IPR007160; DUF362.
DR Pfam: PF04015; DUF362; 1.
DR Pfam: PF00037; fer4; 2.
DR PROSITE: PS00198; 4FE4S_FERREDOXIN; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 377 AA; 40399 MW; 77879A3B2950C3D4 CRC64;

```

```

Query Match 3.3%; Score 7; DB 17; Length 377;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 IDPMKVP 52
DB 248 IDPMKVP 254

RESULT 93
Q8RX25 PRELIMINARY; PRT; 397 AA.
AC Q8RX25;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN AT4G09310.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.

```

```

OX NCBI_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.,
RT "Arabidopsis Full Length cDNA Clones."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Ecker J.R.,
RA Theologis A.,
RT "Arabidopsis Open Reading Frame (ORF) Clones."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY090940; AAM13990.1; -
DR EMBL; AY117341; AAM51416.1; -
DR InterPro; IPR006595; CTLLH.
DR InterPro; IPR006594; LISH.
DR InterPro; IPR003877; SPRY_receptor.
DR Pfam; PF00622; SPRY; 1.
DR PROSITE; PSS0897; CTLLH; 1.
DR PROSITE; PSS0896; LISH; 1.
KW Hypothetical protein.
SQ SEQUENCE 397 AA; 44257 MW; B5C117142814DF70 CRC64;

Query Match 3.3%; Score 7; DB 10; Length 397;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 IDDAIAA 39
DB 265 IDDAIAA 271

RESULT 94
Q88GT2 PRELIMINARY; PRT; 401 AA.
AC
Q88GT2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sulfonate ABC transporter, periplasmic sulfonate-binding protein,
DE putative.
DE
GN PP3636.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzzez A.,
RA Utterback J., Rizso M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Stjepandic D., Hohnselt J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuenmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
```

Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K., Dunn P., Etsu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y., Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L., Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E., Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B., Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P., Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A., Miltschker J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I., Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D., Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M., Sun H., Tallon L.J., Tambunga S., Toriumi M.J., Town C.D., Walker M., Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M., Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.; "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana."; Nature 408:816-820(2000).

Query Match 3.38; Score 7; DB 10; Length 415;
Best Local Similarity 100.0%; Pred.No.2.7e-02; Indels 0; Gaps 0
Matches 7; Conservative 0; Mismatches 0;

QY 139 VVALSLE 145
|||||
DB 173 VVALSLE 179
|||||

RESULT 98
Q8MQF9 PRELIMINARY; PRT; 417 AA.

ID Q8MQF9
AC Q8MQF9;
DT 01-CT-2002 (TrEMBLrel. 22, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-CT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein B0563.6.
OS B0563.6.
GN Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
NCBI_TaxID=6239;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RC MEDLINE=99069613; PubMed=9851916;
RX Wilson R.;
RA "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RC Ravello A.;
RA "The sequence of C. elegans cosmid B0563.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RC Waterston R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR U28740; AA051506.2;
DR WormRep; B0563.6b; CE33513.
CG GO:0016021; C:integral to membrane; IEA.
DR GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO:0007186; P:G-protein coupled receptor protein signalin...; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
PFam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHOODPSN
DR PROSITE; PS00237; G_PROTEIN_RCEP_FL_1; 1.

DR PROSITE; PS0262; G PROTEIN_RECEP_FL_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 417 AA; 47947 MW; 5F085D02F98CF9C2 CRC64;

Query Match 3.3%; Score 7; DB 5; Length 417;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVALSLE 145
|||||
DB 112 VVALSLE 118

RESULT 99
Q9SMS9 PRELIMINARY; PRT; 427 AA.
ID Q9SMS9
AC Q9SMS9
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN T30A10.10 OR AT4G09250.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Van Der Schueren J., Chuang Y.J., Voet M., Robben J.,
RA Volckaert G., Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Robben J., Grymonprez B., Volckaert G., Spiegel L.A., Huang E.N.,
RA Nascimento L.U., de la Bastide M., Vil D.M., Preston R.R., Matero A.,
RA Shah R., O'Shaughnessy A., Rodriguez M., Shekher M., Schutz K.,
RA See L.H., Swaby I., Habermann K., Dedhia N.N., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL117386; CAB55688.1; -.
DR EMBL; AL161514; CAB78048.1; -.
DR PIR; T17123; T17123.
DR InterPro; IPR006594; Lish.
DR InterPro; IPR003877; SPRY_receptor.
DR Pfam; PF00622; SPRY; 1.
DR SMART; SM00667; Lish; 1.
DR PROSITE; PS0896; LISH; 1.
KW Hypothetical protein.
SQ SEQUENCE 427 AA; 47451 MW; 0CB23BD1A35E386C CRC64;

Query Match 3.3%; Score 7; DB 10; Length 427;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 IDDAIAA 39
|||||
DB 343 IDDAIAA 349

RESULT 100
Q84UT3 PRELIMINARY; PRT; 428 AA.
ID Q84UT3
AC Q84UT3
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE F-box.
GN SI-SLFL2.
OS Prunus mume (Japanese flowering apricot).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX NCBI_TaxID=102107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nanko;
RA Entani T., Iwano M., Shiba H., Che F., Isogai A., Takayama S.;
RT "Comparative analysis of the S-locus region of Prunus mume;
RT identification of a pollen-expressed F-box gene with allelic
RT diversity.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB092625; BAC66626.1; -.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR006527; F-box_assoc_1.
DR InterPro; IPR008945; SKP1_Skp2.
DR Pfam; PF00646; F-box; 1.
DR SMART; SM00256; FBOX; 1.
DR TIGRFAMs; TIGR01640; F-box_assoc_1; 1.
SQ SEQUENCE 428 AA; 48563 MW; 6C32BB4063F16752 CRC64;

Query Match 3.3%; Score 7; DB 10; Length 428;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 ALSLEIS 147
|||||
DB 129 ALSLEIS 135

Search completed: August 6, 2004, 16:09:37
JOB time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2004, 15:54:10 ; Search time 54 Seconds

(without alignments)
1114.493 Million cell updates/sec

Title: US-10-024-955-7

Perfect score: 1068

Sequence: 1 MMKFLIIAAVAVFVVSADPI.....VRKEMTKVLAPAFKRELEKN 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A_Geneseq_29Jan04.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001s.*

5: Geneseq2002s.*

6: Geneseq2003as.*

7: Geneseq2003bs.*

8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1068	100.0	213	2	AAR60576 House dus
2	1068	100.0	213	2	AAY25592 D. farina
3	1068	100.0	213	7	ADC34842 House dus
4	962	90.1	215	2	AAY25586 D. pteron
5	962	90.1	215	7	ADC34836 House dus
6	959	89.8	215	2	AAR60575 House dus
7	103	9.6	436	6	ABU44564 Protein e
8	99	9.3	260	4	ABE64121 Drosophil
9	96	9.0	294	5	AAG68268 Human POL
10	96	9.0	294	5	AAG68267 Human POL
11	96	9.0	294	5	AAG68269 Human POL
12	96	9.0	294	6	ASU12102 Novel hum
13	96	9.0	294	6	ASU12101 Novel hum
14	96	9.0	294	6	ASU12100 Novel hum
15	96	9.0	294	7	ADB82715 Human pro
16	96	9.0	294	7	AAB38817 Human syn
17	96	9.0	294	7	AAB38815 Human syn
18	96	9.0	314	4	AAU19614 Human dia
19	96	9.0	314	5	ABP51427 Human MDD
20	93	8.7	436	5	ABP27826 Streptococ
21	91.5	8.6	270	4	AAU66088 Propionib
22	91.5	8.6	270	6	ABM62607 Propionib
23	91.5	8.6	1148	6	ABM64952 Propionib
24	91.5	8.6	1397	6	ABU26950 Protein e
25	91	8.5	294	7	AAB38816 Human syn

26	90.5	8.5	944	6	ABU19882	Abu19882 Protein e
27	90	8.4	1558	3	AAAB18324	AAAB18324 Plasmodi
28	90	8.4	1786	2	AAW24790	AAW24790 P. falcip
29	90	8.4	1787	5	AAU96699	AAU96699 Plasmodi
30	89	8.3	1164	4	ASG25612	Novel hum
31	88.5	8.3	405	3	AAW58737	Novel hum
32	88.5	8.3	458	7	ADE60975	Breast an
33	88.5	8.3	458	7	ADE60975	Human Pro
34	88.5	8.3	458	7	ADE60960	Human Pro
35	88.5	8.3	458	7	ADE60972	Human Pro
36	88.5	8.3	458	7	ADE60966	Human Pro
37	88.5	8.3	458	7	ADE60969	Human Pro
38	88.5	8.3	459	2	AAW54038	Human PK5
39	88.5	8.3	459	6	ABR47454	Human PK5
40	88.5	8.3	459	6	ABR47454	Breast ca
41	88	8.2	436	3	AAW81730	Monoclonal
42	88	8.2	436	3	AAW81730	Streptoco
43	88	8.2	436	6	ABU02185	S. pneumo
44	88	8.2	436	6	ABU46202	Protein e
45	88	8.2	436	6	ABU46202	Protein e
46	87	8.1	436	5	ABP27827	Streptoco
47	87	8.1	436	5	ABU46470	Protein e
48	87	8.1	1640	5	ABB54727	Protein e
49	86.5	8.1	544	6	ABU49623	Protein e
50	86.5	8.1	660	4	AAU33974	Staphyloc
51	86.5	8.1	664	4	AAU36586	Staphyloc
52	86	8.1	280	7	ADB80059	Mycobacte
53	86	8.1	312	4	ABG82219	S. epider
54	85.5	8.0	1396	6	ABU27289	Protein e
55	85.5	8.0	1441	5	AAU75885	Human adh
56	85.5	8.0	2780	4	AAE10924	Mouse m
57	85	8.0	878	5	ABB54705	Lactococc
58	84.5	7.9	293	2	AAW35439	B. burgdo
59	84.5	7.9	436	4	AAU58103	Propionib
60	84.5	7.9	436	6	ABM54622	Propionib
61	84.5	7.9	817	4	AAW81478	S. epider
62	84.5	7.9	817	4	AAW82217	S. epider
63	84.5	7.9	817	6	ABU43148	Protein e
64	84.5	7.9	823	5	ABP39236	Staphyloc
65	84.5	7.9	1396	2	AAW36871	Protein i
66	84.5	7.9	1887	2	ABW58245	Drosophil
67	84	7.9	306	5	ABP27046	Streptoco
68	84	7.9	397	6	ABM72123	Staphyloc
69	84	7.9	410	2	AAW59996	Human end
70	84	7.9	775	5	ABB53885	Lactococc
71	84	7.9	9510	6	AAE36119	Lactococc
72	83.5	7.8	412	4	AAW89772	C. glutani
73	83.5	7.8	517	7	ADC30987	Human nov
74	83.5	7.8	548	6	ABU50581	Protein e
75	83.5	7.8	798	7	ABE15622	Human str
76	83	7.8	583	7	ADC96051	E. faeciu
77	83	7.8	845	5	AAU71861	Leishmani
78	83	7.8	845	5	AAU71861	L. major
79	83	7.8	845	7	ADP78873	Leishmani
80	83	7.8	1060	5	ABB53805	Lactococc
81	82.5	7.7	548	2	AAW16678	Lawsonia
82	82.5	7.7	953	6	AAU11518	Human MDD
83	82.5	7.7	2458	6	ABP59211	Human dru
84	82.5	7.7	2486	4	AAU32848	Novel hum
85	82.5	7.7	2487	6	ABW84649	Human SEC
86	82.5	7.7	2498	5	ABU65149	Human NOV
87	82	7.7	277	4	AAW82447	S. epider
88	82	7.7	357	6	ABW72477	Staphyloc
89	82	7.7	362	5	ABP39212	Streptoco
90	82	7.7	397	2	AAW35476	Chlamydia
91	82	7.7	461	4	AAW35476	C. pneumo
92	82	7.7	461	5	ABW94271	Chlamydia
93	82	7.7	461	6	ABU27059	Protein e
94	82	7.7	487	2	AAW16827	Recombina
95	82	7.7	530	2	AAU04552	Human exp
96	82	7.7	530	2	AAW36857	Protein 1
97	82	7.7	767	4	AAW94022	Human pro
98	82	7.7	767	4	AAW94036	Human pro

99 82 7.7 785 4 AAG91691 Aag91691 C glutami
100 81.5 7.6 227 4 ABB58783 ABB58783 Drosophil

ALIGNMENTS

RESULT 1

AA060576
ID AAR60576 standard; protein; 213 AA.

XX AC AAR60576;

XX DT 25-MAR-2003 (revised)

XX DT 01-APR-1995 (first entry)

XX DE House dust mite allergen DerfVII.

XX KW DerfVII allergen; antiallergic; allergy diagnosis.
XX OS Dermatophagoides farinae.
XX PN WO9420614-A1.
XX PD 15-SEP-1994.
XX PF 11-MAR-1994; 94WO-AU000117.
XX PR 12-MAR-1993; 93US-00031141.
XX PR 22-JUN-1993; 93US-00081540.
XX PA (CHIL-) INST CHILD HEALTH RES.
XX PI Thomas WR, Chua K;
XX DR N-PSDB; AAG71401.
XX PT New nucleic acid encoding specific dust mite allergens - and related
XX PT vectors, transformed cells, peptides and antibodies, useful for
XX PT desensitisation and diagnosis.
XX PS Claim 5; Page 40-41; 67pp; English.
XX CC DerfVII antigen is useful as antiallergic reagent for treating
XX CC sensitivity to house dust mite allergens. (Updated on 25-MAR-2003 to
XX CC correct PN field.)
XX SQ Sequence 213 AA;

Query Match 100.0%; Score 1068; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 5.3e-106;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMKFLIIAAVAVFVAVSADPIHYDKITEINKAIDDAIAAIEQSETIDPMKVPDADKFER 60
DB 1 MMKFLIIAAVAVFVAVSADPIHYDKITEINKAIDDAIAAIEQSETIDPMKVPDADKFER 60
QY 61 HVGIVDFKGLAMRNIEARGLKQKQGDANVKGEIGVKAHLIGVHDDIVSMEYDLAY 120
DB 61 HVGIVDFKGLAMRNIEARGLKQKQGDANVKGEIGVKAHLIGVHDDIVSMEYDLAY 120
QY 121 KLGLDHPHTHVISDIDQDFVWALSLEISDEGNITMTSFEVRFANVNVHIGLSILDPIFG 180
DB 121 KLGLDHPHTHVISDIDQDFVWALSLEISDEGNITMTSFEVRFANVNVHIGLSILDPIFG 180
QY 181 VLSVDLTAIFQDVTVRKEMTKVLAPAFKRELEKN 213
DB 181 VLSVDLTAIFQDVTVRKEMTKVLAPAFKRELEKN 213

RESULT 2

AA25592

ID AAY25592 standard; protein; 213 AA.
XX AC AAY25592;
XX DT 30-SEP-1999 (first entry)
XX DE D. farinae allergen Der f 7 protein fragment.
XX KW Major histocompatibility complex; class II; desensitising; human;
XX KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; stinging;
XX KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
XX KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
XX KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
XX KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity.
XX OS Dermatophagoides farinae.

XX PN WO934826-A1.
XX PD 15-JUL-1999.
XX PF 11-JAN-1999; 99WO-GB000080.
XX PR 09-JAN-1998; 98GB-00000445.
XX PR 21-SEP-1998; 98GB-00020474.
XX PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX PI Larche M, Kay AB;
XX DR WPI; 1999-458255/38.
XX PT Desensitizing patients to polypeptide allergens.
XX PS Example 6; Page 52; 117pp; English.
XX CC This invention describes a novel method of desensitizing a patient to a
XX CC polypeptide allergen and comprises administering to the patient a peptide
XX CC derived from the allergen where restriction to a MHC Class II molecule
XX CC possessed by the patient can be demonstrated for the peptide and the
XX CC peptide is able to induce a late phase response in an individual who
XX CC possesses the MHC Class II molecule. The methods can be used for
XX CC desensitising patients to allergens present in e.g. grass, tree and weed
XX CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
XX CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
XX CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
XX CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
XX CC Tenibrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
XX CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
XX CC produce immunological vaccines which may be used to prevent and/or treat
XX CC conditions involving hypersensitivity to allergens. This sequence
XX CC represents the Dermatophagoides farinae allergen Der f 7
XX SQ Sequence 213 AA;

Query Match 100.0%; Score 1068; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 5.3e-106;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMKFLIIAAVAVFVAVSADPIHYDKITEINKAIDDAIAAIEQSETIDPMKVPDADKFER 60
DB 1 MMKFLIIAAVAVFVAVSADPIHYDKITEINKAIDDAIAAIEQSETIDPMKVPDADKFER 60
QY 61 HVGIVDFKGLAMRNIEARGLKQKQGDANVKGEIGVKAHLIGVHDDIVSMEYDLAY 120
DB 61 HVGIVDFKGLAMRNIEARGLKQKQGDANVKGEIGVKAHLIGVHDDIVSMEYDLAY 120
QY 121 KLGLDHPHTHVISDIDQDFVWALSLEISDEGNITMTSFEVRFANVNVHIGLSILDPIFG 180
DB 121 KLGLDHPHTHVISDIDQDFVWALSLEISDEGNITMTSFEVRFANVNVHIGLSILDPIFG 180
QY 181 VLSVDLTAIFQDVTVRKEMTKVLAPAFKRELEKN 213

Db 181 VLSVLTALFQDTRKEMTKVLAPAFKRELEKN 213

RESULT 3
ADC34842
ID ADC34842 standard; protein; 213 AA.

XX AC ADC34842;
XX AC ADC34842;
DT 18-DEC-2003 (first entry)

XX DE House dust mite allergen Der f 7.
XX KW house dust mite; allergen; antigen; hyporesponsive; desensitisation;
XX KW immunomodulator; gene therapy.

XX OS Dermatophagoides farinae.
XX FN W02003047618-A2.

XX PD 12-JUN-2003.

XX PF 03-DEC-2002; 2002WO-GB005548.

XX PR 05-DEC-2001; 2001US-0338385P.

XX PA (CIRC-) CIRCASSIA LTD.

XX PI Larche M, Ledger PW;

XX DR WPI; 2003-523267/49.

XX PT Desensitizing an individual to a selected polypeptide antigen comprises
XX PT administering a composition containing polypeptide antigens in an amount
XX PT that generates a state of hyporesponsiveness to the antigen to allow
XX PT desensitization.

XX PS Disclosure; Page 23; 57pp; English.

XX CC The invention relates to a novel method for desensitising an individual
XX CC to a selected polypeptide antigen. The method comprises administering a
XX CC composition that contains polypeptide antigens in an amount that
XX CC generates in the individual a state of hyporesponsiveness to the antigen
XX CC to allow desensitisation to one or more polypeptide antigens. The method
XX CC of the invention has immunomodulator activity, and may have a use in gene
XX CC therapy. The composition and method are useful in manufacturing a
XX CC medicament for desensitising an individual to a selected polypeptide
XX CC antigen or for generating in the individual a state of hyporesponsiveness
XX CC to the antigen to allow desensitisation to one or more polypeptide
XX CC antigens. The present sequence is used in the exemplification of the
XX CC invention.

XX SQ Sequence 213 AA;

Query Match 100.0%; Score 1068; DB 7; Length 213;
Best Local Similarity 100.0%; Pred. No. 5.3e-106;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMKLLIAAVAFVAVSADPIHYDKITEINKAIDDAIAAEQSETIDPMKVPDHDKFER 60
Db 1 MMKLLIAAVAFVAVSADPIHYDKITEINKAIDDAIAAEQSETIDPMKVPDHDKFER 60
Qy 61 HVGIVDFKGLAMRNIEARGLKQMKRQGDANVKGEIGVKAHLIJGVHDDIVSMEYDLAY 120
Db 61 HVGIVDFKGLAMRNIEARGLKQMKRQGDANVKGEIGVKAHLIJGVHDDIVSMEYDLAY 120
Qy 121 KLGLDHPHTTHVISDIQDFVWALSLEISDEGNITMTSFEVRQFANVNVHIGLSILDPIFG 180
Db 121 KLGLDHPHTTHVISDIQDFVWALSLEISDEGNITMTSFEVRQFANVNVHIGLSILDPIFG 180
Qy 181 VLSVLTALFQDTRKEMTKVLAPAFKRELEKN 213
Db 181 VLSVLTALFQDTRKEMTKVLAPAFKRELEKN 213

RESULT 4

AY25586
ID AAY25586 standard; protein; 215 AA.

XX AC AAY25586;

XX DT 30-SEP-1999 (first entry)

XX DE D. pteronyssinus allergen Der p 7 protein fragment.

XX KW Major histocompatibility complex; class II; desensitising; human;
XX KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
XX KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
XX KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
XX KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
XX KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity.

XX OS Dermatophagoides pteronyssinus.

XX FN W09934826-A1.

XX PD 15-JUL-1999.

XX PF 11-JAN-1999; 99WO-GB000080.

XX PR 09-JAN-1998; 98GB-00000445.

XX PR 21-SEP-1998; 98GB-00020474.

XX PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

XX PI Larche M, Kay AB;

XX DR WPI; 1999-458255/38.

XX PT Desensitizing patients to polypeptide allergens.

XX PS Example 6; Page 51; 117pp; English.

XX CC This invention describes a novel method of desensitising a patient to a
XX CC polypeptide allergen and comprises administering to the patient a peptide
XX CC derived from the allergen where restriction to a MHC Class II molecule
XX CC possessed by the patient can be demonstrated for the peptide and the
XX CC peptide is able to induce a late phase response in an individual who
XX CC possesses the MHC Class II molecule. The methods can be used for
XX CC desensitising patients to allergens present in e.g. grass, tree and weed
XX CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
XX CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
XX CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
XX CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
XX CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
XX CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
XX CC produce immunological vaccines which may be used to prevent and/or treat
XX CC conditions involving hypersensitivity to allergens. This sequence
XX CC represents the house dust mite (Dermatophagoides pteronyssinus) allergen
XX CC Der p 7

XX SQ Sequence 215 AA;

Query Match 90.1%; Score 962; DB 2; Length 215;
Best Local Similarity 85.9%; Pred. No. 1.3e-94;
Matches 183; Conservative 18; Mismatches 12; Indels 0; Gaps 0;

Qy 1 MMKLLIAAVAFVAVSADPIHYDKITEINKAIDDAIAAEQSETIDPMKVPDHDKFER 60
Db 1 MMKLLIAAVAFVAVSADPIHYDKITEINKAIDDAIAAEQSETIDPMKVPDHDKFER 60
Qy 61 HVGIVDFKGLAMRNIEARGLKQMKRQGDANVKGEIGVKAHLIJGVHDDIVSMEYDLAY 120
Db 61 HVGIVDFKGLAMRNIEARGLKQMKRQGDANVKGEIGVKAHLIJGVHDDIVSMEYDLAY 120
Qy 121 KLGLDHPHTTHVISDIQDFVWALSLEISDEGNITMTSFEVRQFANVNVHIGLSILDPIFG 180

Db 121 KLGDLPNTHVTSIDQDFVVELSLEVEGNTLTSTFEVRFQFANVNVNHIGLSILDPIFA 180
QY 181 VLSVDLTAIFQDVTVRKEMTKVLAPAFKKELEKN 213
Db 181 VLSVDLTAIFQDVTVRKEMTKVLAPAFKKELEKN 213

RESULT 5
ADC34836
ID ADC34836 standard; protein; 215 AA.
XX AC ADC34836;
XX DT 18-DEC-2003 (first entry)
XX DE House dust mite allergen Der p 7.
XX KW house dust mite; allergen; antigen; hyporesponsive; desensitization;
XX KW immunomodulator; gene therapy.
XX OS Dermatophagoides pteronyssinus.
XX PN WO2003047618-A2.
XX PD 12-JUN-2003.
XX PF 05-DEC-2002; 2002WO-GB005548.
XX PR 05-DEC-2001; 2001US-0338385P.
XX PA (CIRC-) CIRCASSIA LTD.
XX PI Larche M, Ledger PW;
XX WPI; 2003-523267/49.
XX PT Desensitizing an individual to a selected polypeptide antigen comprises
PT administering a composition containing polypeptide antigens in an amount
PT that generates a state of hyporesponsiveness to the antigen to allow
PT desensitization.
XX PS Disclosure; Page 21; 57pp; English.
XX CC The invention relates to a novel method for desensitizing an individual
CC to a selected polypeptide antigen. The method comprises administering a
CC composition that contains polypeptide antigens in an amount that
CC generates in the individual a state of hyporesponsiveness to the antigen
CC to allow desensitization to one or more polypeptide antigens. The method
CC of the invention has immunomodulator activity, and may have a use in gene
CC therapy. The composition and method are useful in manufacturing a
CC medicament for desensitizing an individual to a selected polypeptide
CC antigen or for generating in the individual a state of hyporesponsiveness
CC to the antigen to allow desensitization to one or more polypeptide
CC antigens. The present sequence is used in the exemplification of the
CC invention.
XX SQ Sequence 215 AA;
Query Match 90.1%; Score 962; DB 7; Length 215;
Best Local Similarity 85.9%; Pred. No. 1.3e-94;
Matches 183; Conservative 18; Mismatches 12; Indels 0; Gaps 0;
QY 1 MKKELLIAAFVAVSADPIHYDKITEINKAIDATAAEQSETIDPMKVPDHADKFER 60
Db 1 MKKLLLTAAAFVAVSADPIHYDKITEINKAIDATAAEQSETIDPMKVPDHADKFER 60
QY 61 HVGIVDFKGEIAMRNIEARGIKQMKRGDANVKGEIGVKAHLIGVHDDIVSMYDLAY 120
Db 61 HIGIIDLKGLDMRNIQVRGLKQMKRGDANVKSEGDGVKAHLIGVHDDIVSMYDLAY 120
QY 121 KLGDLPNTHVTSIDQDFVVELSLEVEGNTLTSTFEVRFQFANVNVNHIGLSILDPIFG 180

Db 121 KLGDLPNTHVTSIDQDFVVELSLEVEGNTLTSTFEVRFQFANVNVNHIGLSILDPIFA 180
QY 181 VLSVDLTAIFQDVTVRKEMTKVLAPAFKKELEKN 213
Db 181 VLSVDLTAIFQDVTVRKEMTKVLAPAFKKELEKN 213

RESULT 6
AAR60575
ID AAR60575 standard; protein; 215 AA.
XX AC AAR60575;
XX DT 25-MAR-2003 (revised)
XX DT 01-APR-1995 (first entry)
XX DE House dust mite allergen DerpVII cDNA.
XX KW DerpVII allergen; anti-allergic; allergy diagnosis.
XX OS Dermatophagoides pteronyssinus.
XX PN WO9420614-A1.
XX PD 15-SEP-1994.
XX PF 11-MAR-1994; 94WO-AU0000117.
XX PR 12-MAR-1993; 93US-00031141.
XX PR 22-JUN-1993; 93US-00081540.
XX PA (CHIL-) INST CHILD HEALTH RES.
XX PI Thomas WR, Chua K;
XX WPI; 1994-303021/37.
XX N-PSDB; AAQ71400.
XX PT New nucleic acid encoding specific dust mite allergens - and related
PT vectors, transformed cells, peptides and antibodies, useful for
PT desensitization and diagnosis.
XX PS Claim 7; Page 36-37; 67pp; English.
XX CC DerpVII antigen is useful as anti-allergic reagent for treating
CC sensitivity to house dust mite allergens. The DNA can be used as a probe
CC to detect the sensitivity of an individual to the allergen. (Updated on 25
CC -MAR-2003 to correct PN field.)
XX SQ Sequence 215 AA;
Query Match 89.8%; Score 959; DB 2; Length 215;
Best Local Similarity 85.4%; Pred. No. 2.6e-94;
Matches 182; Conservative 19; Mismatches 12; Indels 0; Gaps 0;
QY 1 MKKELLIAAFVAVSADPIHYDKITEINKAIDATAAEQSETIDPMKVPDHADKFER 60
Db 1 MKKLLLTAAAFVAVSADPIHYDKITEINKAIDATAAEQSETIDPMKVPDHADKFER 60
QY 61 HVGIVDFKGEIAMRNIEARGIKQMKRGDANVKGEIGVKAHLIGVHDDIVSMYDLAY 120
Db 61 HIGIIDLKGLDMRNIQVRGLKQMKRGDANVKSEGDGVKAHLIGVHDDIVSMYDLAY 120
QY 121 KLGDLPNTHVTSIDQDFVVELSLEVEGNTLTSTFEVRFQFANVNVNHIGLSILDPIFG 180
Db 121 KLGDLPNTHVTSIDQDFVVELSLEVEGNTLTSTFEVRFQFANVNVNHIGLSILDPIFA 180
QY 181 VLSVDLTAIFQDVTVRKEMTKVLAPAFKKELEKN 213
Db 181 VLSVDLTAIFQDVTVRKEMTKVLAPAFKKELEKN 213

RESULT 7

OS Homo sapiens.
PN WO200179294-A2.
XX 25-OCT-2001.
PD 19-APR-2001; 2001WO-US012854.
XX 19-APR-2000; 2000US-0198293P.
XX 20-APR-2000; 2000US-0198645P.
PR 25-APR-2000; 2000US-0199476P.
PR 26-APR-2000; 2000US-0199880P.
PR 26-APR-2000; 2000US-0200024P.
PR 26-APR-2000; 2000US-0200025P.
PR 09-JUN-2000; 2000US-0210809P.
PR 17-JUL-2000; 2000US-0218591P.
PR 11-AUG-2000; 2000US-0224610P.
PR 09-FEB-2001; 2001US-0267673P.
PR 27-FEB-2001; 2001US-0271814P.
XX (CURA-) CURAGEN CORP.
PA Taupier RJ, Vernet CAM, Fernandes E, Shinkets RA, Majumder K;
PI Padigaru M, Colman SD, Zerhusen BD, Spytek KA, Burgess CE, Liu X;
XX WPI; 2002-017601/02.
XX N-PSDB; ABA03883.
XX New isolated polypeptides for treating a broad range of pathological
PT states, e.g., depression, stroke, Parkinson's disease, Huntington's
PT disease, Tourette's syndrome, amyotrophic lateral sclerosis, head trauma,
PT and Alzheimer's.
XX Claim 1; Page 57; 155pp; English.
XX The present invention describes polypeptides (I), designated POLYX
XX polypeptides (e.g. POLY1, POLY2, etc.), and the polynucleotide sequences
CC (II) encoding them. POLY1-4 are members of the gamma aminobutyric acid
CC (GABA) receptor family; POLY5-8 are members of the complement receptor
CC factor (EGF) family; POLY9-11 are members of the haematopoietic stem and progenitor cell
CC family; POLY12 is a member of the sulphotransferase family; POLY14
CC (HSPC) family; POLY13 is a member of the syntaxin family; and POLY17
CC -16 are members of the syntaxin family; and POLY17 is a member of the
CC prohibitin family. (I) and (II) can have antidepressant,
CC neuroprotective, antiparkinsonian, nootropic, relaxant, anticonvulsant,
CC neuroleptic, neuroprotective, antialcoholic, cardiant, tranquiliser and
CC antiarrhythmic activities. (I) and (II) can be used for treating or
CC preventing a POLYX-associated disorder in humans as a therapeutic in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from a POLYX-associated disorder, for treating a
CC pathological state in a mammal, especially patients suffering from, e.g.,
CC psychiatric and medical conditions, depression, stroke, Parkinson's
CC disease, Huntington's disease, Tourette's syndrome, amyotrophic lateral
CC sclerosis, head trauma, Alzheimer's disease, alcoholism, vigilance,
CC anxiety, muscle tension, epileptogenic activity and memory functions,
CC cardiomyopathy and arrhythmic right ventricular dysplasia. The
CC present sequence represents POLY14, which was identified on chromosome 1
XX Sequence 294 AA;
SQ Query Match
Best Local Similarity 9.0%; Score 96; DB 5; Length 294;
Matches 44; Conservative 31; Mismatches 58; Indels 52; Gaps 10;
Qy 27 BEINKAIDDAIAEQSETID-PMKV-----PHADKFERHVGIVDF-----KGE- 70
Db 103 EYINRSLNDLVKEVKKSEVENGPSSVTRILKSQAAMP-RHFQOIFNFIYNDITAAKQEK 161
Qy 71 ---LAMENIARG-----LQMKKRGDANVKCEE-----GVKALH---LIGVHDIVS 113
Db 162 KCTILKQLEVAGKEMSEEDVNDMLHQQKVEFNESLLTINTKALQSIEQRHKLVN 221
Qy 114 MEYDLAYKLGDLHPTTHTVISDIQDFVVALSLE-SDEG-----NITMTSFEVQRQFANVNH 169

Db 222 LE-----NQIKLRDLFIQISLVVEEQSSINNIEVTNVSTKEYVNTKKE 267
Qy 170 GGLSI 174
Db 268 FGLAV 272
RESULT 11
AAG68269
ID AAG68269 standard; protein; 294 AA.
XX AAG68269;
AC AAG68269;
XX 14-FEB-2002 (first entry)
XX Human POLY16 protein sequence SEQ ID NO:32.
XX Human; POLYX; gamma aminobutyric acid receptor; GABA receptor;
KW epidermal growth factor; EGF; complement receptor; HSPC; syntaxin;
KW haematopoietic stem and progenitor cell; sulphotransferase; prohibitin;
KW antidepressant; neuroprotective; antiparkinsonian; nootropic; relaxant;
KW anticonvulsant; neuroleptic; neuroprotective; antialcoholic; cardiant;
KW tranquiliser; antiarrhythmic; psychiatric; medical; depression; stroke;
KW Parkinson's disease; Huntington's disease; Tourette's syndrome; anxiety;
KW amyotrophic lateral sclerosis; head trauma; Alzheimer's disease;
KW alcoholism; vigilance; muscle tension; epileptogenic; memory function;
KW cardiomyopathy; arrhythmic right ventricular dysplasia.
XX Homo sapiens.
OS Homo sapiens.
XX WO200179294-A2.
XX 25-OCT-2001.
XX 19-APR-2001; 2001WO-US012854.
XX 19-APR-2000; 2000US-0198293P.
XX 20-APR-2000; 2000US-0198645P.
PR 25-APR-2000; 2000US-0199476P.
PR 26-APR-2000; 2000US-0199880P.
PR 26-APR-2000; 2000US-0200024P.
PR 26-APR-2000; 2000US-0200025P.
PR 09-JUN-2000; 2000US-0210809P.
PR 17-JUL-2000; 2000US-0218591P.
PR 11-AUG-2000; 2000US-0224610P.
PR 09-FEB-2001; 2001US-0267673P.
PR 27-FEB-2001; 2001US-0271814P.
XX (CURA-) CURAGEN CORP.
PA Taupier RJ, Vernet CAM, Fernandes E, Shinkets RA, Majumder K;
PI Padigaru M, Colman SD, Zerhusen BD, Spytek KA, Burgess CE, Liu X;
XX WPI; 2002-017601/02.
XX N-PSDB; ABA03885.
XX New isolated polypeptides for treating a broad range of pathological
PT states, e.g., depression, stroke, Parkinson's disease, Huntington's
PT disease, Tourette's syndrome, amyotrophic lateral sclerosis, head trauma,
PT and Alzheimer's.
XX Claim 1; Page 61; 155pp; English.
XX The present invention describes polypeptides (I), designated POLYX
CC polypeptides (e.g. POLY1, POLY2, etc.), and the polynucleotide sequences
CC (II) encoding them. POLY1-4 are members of the gamma aminobutyric acid
CC (GABA) receptor family; POLY5-8 are members of the epidermal growth
CC factor (EGF) family; POLY9-11 are members of the complement growth
CC factor (EGF) family; POLY9-11 are members of the complement receptor
CC family; POLY12 is a member of the haematopoietic stem and progenitor cell
CC (HSPC) family; POLY13 is a member of the syntaxin family; and POLY17
CC -16 are members of the syntaxin family; and POLY17 is a member of the
CC prohibitin family. (I) and (II) can have antidepressant,

CC cerebroprotective, antiparkinsonian, nootropic, relaxant, anticonvulsant,
 CC neuroleptic, neuroprotective, antialcoholic, cardiant, tranquiliser and
 CC antiarhythmic activities. (I) and (II) can be used for treating or
 CC preventing a POLYX-associated disorder in humans as a therapeutic in the
 CC manufacture of a medicament for treating a syndrome associated with a
 CC human disease selected from a POLYX-associated disorder, for treating a
 CC pathological state in a mammal, especially patients suffering from, e.g.,
 CC psychiatric and medical conditions, depression, stroke, Parkinson's
 CC disease, Huntington's disease, Tourette's syndrome, amyotrophic lateral
 CC sclerosis, head trauma, Alzheimer's disease, alcoholism, vigilance,
 CC anxiety, muscle tension, epileptogenic activity and memory functions,
 CC cardiomyopathy and arrhythmogenic right ventricular dysplasia. The
 CC present sequence represents POLY16
 XX
 SQ Sequence 294 AA;

Query Match 9.0%; Score 96; DB 5; Length 294;
 Best Local Similarity 23.8%; Pred. No. 0.15; Indels 52; Gaps 10;
 Matches 44; Conservative 31; Mismatches 58; Indels 52; Gaps 10;
 QY 27 EEINKAIDDAIAAEQSETID-PMKV-----PDHAKFERHVGIVDF-----KGE- 70
 DB 103 EYINRSLNDLVKEVKSEVNGPSSVTRILKSQAAMF-RHFQOIMFYNDTIAAKQEK 161
 QY 71 ---LAVRNIARG-----LKQMKRGDANVKGE-----GIVKAHL--LIGVHDDIVS 113
 DB 162 CKTFILRQLEVAGKEMSEEDVNDMLHQKWEVFNESLLEINITYKAQSEIEQRHKLNV 221
 QY 114 MEYDLAYKLGDLHPHTTHVISDIDQDFVALSLSEISDEG-----NITWTSFEVRQFANVNH 169
 DB 222 LE-----NQIKDLRDLFIQISLLEVEGQGSINNIENTVNSTKEYVNTKEK 267
 QY 170 GGLSI 174
 DB 268 FGLAV 272

RESULT 12
 ABU12102
 ID ABU12102 standard; protein; 294 AA.

XX AC ABU12102;
 XX DT 17-FEB-2003 (first entry)
 XX DE Novel human syntaxin-like protein #3.

KW Gamma-aminobutyric acid receptor-like protein; depression; stroke;
 KW GABA receptor-like protein; Parkinson's disease; Huntington's disease;
 KW Tourette's syndrome; amyotrophic lateral sclerosis; head trauma;
 KW Alzheimer's disease; alcoholism; vigilance; anxiety; muscle tension;
 KW epileptogenic activity; memory; cardiomyopathy; cancer; angiogenesis;
 KW arrhythmogenic right ventricular dysplasia; renal disease; diabetes;
 KW Epidermal growth factor like protein; leukaemia; lupus; anaemia; ulcer;
 KW haematopoietic stem and progenitor cell like protein; cirrhosis;
 KW sulfolipase-like protein; cholangitis; hepatitis; hyperthyroidism;
 KW developmental disorder; Syntaxin-like protein; myxoid liposarcoma;
 KW asthma; Lambert-Baton myasthenic syndrome; acute myeloid leukaemia;
 KW transgenic animal.

XX Homo sapiens.

XX OS US2002123612-A1.

XX PN 05-SEP-2002.

XX PD 03-JUL-2001; 2001US-00898570.

XX PF 19-APR-2000; 2000US-0198293P.

XX PR 20-APR-2000; 2000US-0198645P.

XX PR 25-APR-2000; 2000US-0199476P.

XX PR 26-APR-2000; 2000US-0199880P.

XX PR 26-APR-2000; 2000US-0200024P.

PR 26-APR-2000; 2000US-0200025P.
 PR 09-JUN-2000; 2000US-0210809P.
 PR 03-JUL-2000; 2000US-0215855P.
 PR 17-JUL-2000; 2000US-0218591P.
 PR 11-AUG-2000; 2000US-0224610P.
 PR 27-FEB-2001; 2001US-0271814P.
 XX (GERL/) GERLACH V L.
 PA (ELLE/) ELLERMAN K.
 PA (MACD/) MACDOUGALL J R.
 PA (SMIT/) SMITHSON G.
 XX Gerlach VL, Ellerman K, Macdougall JR, Smithson G;
 PI WPI; 2003-066815/06.
 DR N-PSDB; ABX56481.
 DR Novel polypeptides and nucleic acids which are members of epidermal
 PT growth factor complement receptor families for diagnosing and treating
 PT psychiatric conditions, depression, stroke, Alzheimer's and Parkinson's
 PT disease.
 XX
 PS Claim 1; Page 41; 91pp; English.

CC The invention describes an isolated POLYX (POLY1-17) polypeptide and its
 CC variant. POLYX polypeptides (especially POLY5, POLY6 and POLY7), the
 CC polynucleotides encoding them (I) and an anti-POLYX-antibody (III) are
 CC useful for treating or preventing a pathology associated with human
 CC polypeptide in humans and for treating a syndrome associated with human
 CC disease. POLYX polypeptide is also useful for identifying an agent that
 CC binds to POLYX and a cell expressing POLYX is useful for identifying a
 CC therapeutic agent for use in treatment of a pathology related to aberrant
 CC expression or physiological interactions of the polypeptide. (III) is
 CC useful for treating a pathological state in a mammal and for determining
 CC the presence or amount of POLYX in a sample. POLY1-4 (GABA receptor-like
 CC proteins) are useful for the treatment of psychiatric and medical
 CC conditions, depression, stroke, Parkinson's disease, Huntington's
 CC disease, Tourette's syndrome, amyotrophic lateral sclerosis, head trauma,
 CC Alzheimer's disease, alcoholism, vigilance, anxiety, muscle tension,
 CC epileptogenic activity and memory functions, cardiomyopathy and
 CC arrhythmogenic right ventricular dysplasia. POLY5-8 (Epidermal growth
 CC factor like proteins) may be useful for treating cancer, aberrant
 CC angiogenesis, renal disease and diabetes. POLY12 (haematopoietic stem and
 CC progenitor cell like protein) may be useful for treatment of leukaemia,
 CC lupus and anaemia. POLY13 (sulfolipase-like protein) may be useful
 CC for treating cirrhosis, cholangitis, hepatitis, ulcers, hyperthyroidism
 CC and developmental disorders. POLY14-16 (Syntaxin-like proteins) may be
 CC useful in treatment of Lambert-Eaton myasthenic syndrome, asthma, myxoid
 CC liposarcoma and acute myeloid leukaemia, and POLY18 may be useful in
 CC treatment of cancers. Cells comprising (I) are useful for producing non-
 CC human transgenic animals which are useful for studying the function
 CC and/or activity of POLYX protein and for identifying and/or evaluating
 CC modulators of POLYX protein activity. This is the amino acid sequence of
 CC a novel human protein
 XX
 SQ Sequence 294 AA;

Query Match 9.0%; Score 96; DB 6; Length 294;
 Best Local Similarity 23.8%; Pred. No. 0.15;
 Matches 44; Conservative 31; Mismatches 58; Indels 52; Gaps 10;

QY 27 EEINKAIDDAIAAEQSETID-PMKV-----PDHAKFERHVGIVDF-----KGE- 70
 DB 103 EYINRSLNDLVKEVKSEVNGPSSVTRILKSQAAMF-RHFQOIMFYNDTIAAKQEK 161
 QY 71 ---LAVRNIARG-----LKQMKRGDANVKGE-----GIVKAHL--LIGVHDDIVS 113
 DB 162 CKTFILRQLEVAGKEMSEEDVNDMLHQKWEVFNESLLEINITYKAQSEIEQRHKLNV 221
 QY 114 MEYDLAYKLGDLHPHTTHVISDIDQDFVALSLSEISDEG-----NITWTSFEVRQFANVNH 169
 DB 222 LE-----NQIKDLRDLFIQISLLEVEGQGSINNIENTVNSTKEYVNTKEK 267

QY 170 GGLSI 174
Db 268 FGLAV 272

RESULT 13
ABU12101
ID ABU12101 standard; protein; 294 AA.
AC ABU12101;
XX
XX 17-FEB-2003 (first entry)
DE Novel human syntaxin-like protein #2.
XX
XX Gamma-aminobutyric acid receptor-like protein; depression; stroke;
KW GABA receptor-like protein; Parkinson's disease; Huntington's disease;
KW Tourette's syndrome; amyotrophic lateral sclerosis; head trauma;
KW Alzheimer's disease; alcoholism; vigilance; anxiety; muscle tension;
KW epileptogenic activity; memory; cardiomyopathy; cancer; angiogenesis;
KW arrhythmogenic right ventricular dysplasia; renal disease; diabetes;
KW haematopoietic stem and progenitor cell like protein; cirrhosis;
KW epidermal growth factor like protein; leukaemia; lupus; anaemia; ulcer;
KW sulfotransferase-like protein; cholangitis; hepatitis; hyperthyroidism;
KW developmental disorder; Syntaxin-like protein; myxoid liposarcoma;
KW asthma; Lambert-Eaton myasthenic syndrome; acute myeloid leukaemia;
KW transgenic animal.
XX
XX Homo sapiens.
XX
XX US2002123612-A1.
XX
XX 05-SEP-2002.
XX
XX 03-JUL-2001; 2001US-00898570.
XX
XX 19-APR-2000; 2000US-0198293P.
XX 20-APR-2000; 2000US-0198645P.
XX 25-APR-2000; 2000US-0199476P.
XX 26-APR-2000; 2000US-0199880P.
XX 26-APR-2000; 2000US-0200024P.
XX 26-APR-2000; 2000US-0200025P.
XX 09-JUN-2000; 2000US-0210809P.
XX 03-JUL-2000; 2000US-0215855P.
XX 17-JUL-2000; 2000US-0218591P.
XX 11-AUG-2000; 2000US-0224610P.
XX 27-FEB-2001; 2001US-0271814P.
XX
XX (GERL/) GERLACH V L.
XX (ELLE/) ELLERMAN K.
XX (MACD/) MACDOUGALL J R.
XX (SMIT/) SMITHSON G.
XX
XX Gerlach VL, Ellerman K, Macdougall JR, Smithson G;
XX
XX WPI; 2003-086815/06.
XX N-PSDB; ABX56480.
XX
XX Novel polypeptides and nucleic acids which are members of epidermal
PT growth factor, complement receptor families for diagnosing and treating
PT psychiatric conditions, depression, stroke, Alzheimer's and Parkinson's
PT disease.
XX
XX Claim 1; Page 30; 91pp; English.

The invention describes an isolated POLYX (POLY1-17) polypeptide and its variant. POLYX polypeptides (especially POLY5, POLY6 and POLY7), the polynucleotides encoding them (I) and an anti-POLYX-antibody (III) are useful for treating or preventing a pathology associated with POLYX polypeptide in humans and for treating a syndrome associated with human disease. POLYX polypeptide is also useful for identifying an agent that binds to POLYX and a cell expressing POLYX is useful for identifying a therapeutic agent for use in treatment of a pathology related to aberrant

expression or physiological interactions of the polypeptide. (III) is useful for treating a pathological state in a mammal and for determining the presence or amount of POLYX in a sample. POLY1-4 (GABA receptor-like proteins) are useful for the treatment of psychiatric and medical conditions, depression, stroke, Parkinson's disease, Huntington's disease, Tourette's syndrome, amyotrophic lateral sclerosis, head trauma, Alzheimer's disease, alcoholism, vigilance, anxiety, muscle tension, epileptogenic activity and memory functions, cardiomyopathy and arrhythmogenic right ventricular dysplasia. POLY5-8 (epidermal growth factor like proteins) may be useful for treating cancer, aberrant angiogenesis, renal disease and diabetes. POLY12 (haematopoietic stem and progenitor cell like protein) may be useful for treatment of leukaemia, lupus and anaemia. POLY13 (sulfotransferase-like protein) may be useful for treating cirrhosis, cholangitis, hepatitis, ulcers, hyperthyroidism and developmental disorders. POLY14-16 (Syntaxin-like proteins) may be useful in treatment of Lambert-Eaton myasthenic syndrome, asthma, myxoid liposarcoma and acute myeloid leukaemia, and POLY 18 may be useful in treatment of cancers. Cells comprising (I) are useful for producing non-human transgenic animals which are useful for studying the function and/or activity of POLYX protein and for identifying and/or evaluating modulators of POLYX protein activity. This is the amino acid sequence of a novel human protein

Sequence 294 AA;
Query Match 9.0%; Score 96; DB 6; Length 294;
Best Local Similarity 23.8%; Pred. No. 0.15;
Matches 44; Conservative 31; Mismatches 58; Indels 52; Gaps 10;
QY 27 EEINKAIDDAIAIEQSETID-PMKV-----PHADKFEHVGIVDF-----KGE- 70
Db 103 EVINRSLNDLVKVKSEVNGSPVVTRILKSQAAMF-RHQIIMPIYNDTIAAKQEK 161
QY 71 ---LAMRNIEARG-----LQMKKRGDANVKGE-----GIVKAHL--LIGVHDDIVS 113
Db 162 CKTFILROLEVAGKEMSEEDVNDMLHQKWEVFNESLLEINITYKAQLSEIQHKELVN 221
QY 114 MEYDLAYKLGDLHPHTTHVISDIQFVVALSLEISDEG-----NITMTGFVQCFANVNH 169
Db 222 LE-----NQIKDLRLDFIQISLLVEEGESINNIENTVNSTKEYVNTTKEK 267
QY 170 GGLSI 174
Db 268 FGLAV 272

RESULT 14
ABU12100
ID ABU12100 standard; protein; 294 AA.
XX
XX ABU12100;
XX
XX 17-FEB-2003 (first entry)
XX
XX Novel human syntaxin-like protein #1.
XX
XX Gamma-aminobutyric acid receptor-like protein; depression; stroke;
KW GABA receptor-like protein; Parkinson's disease; Huntington's disease;
KW Tourette's syndrome; amyotrophic lateral sclerosis; head trauma;
KW Alzheimer's disease; alcoholism; vigilance; anxiety; muscle tension;
KW epileptogenic activity; memory; cardiomyopathy; cancer; angiogenesis;
KW arrhythmogenic right ventricular dysplasia; renal disease; diabetes;
KW epidermal growth factor like protein; leukaemia; lupus; anaemia; ulcer;
KW haematopoietic stem and progenitor cell like protein; cirrhosis;
KW sulfotransferase-like protein; cholangitis; hepatitis; hyperthyroidism;
KW developmental disorder; Syntaxin-like protein; myxoid liposarcoma;
KW asthma; Lambert-Eaton myasthenic syndrome; acute myeloid leukaemia;
KW transgenic animal.
XX
XX Homo sapiens.
XX
XX US2002123612-A1.
XX

PD 05-SEP-2002.
XX 03-JUL-2001; 2001US-00898570.
XX 19-APR-2000; 2000US-0198293P.
PR 20-APR-2000; 2000US-0198645P.
PR 25-APR-2000; 2000US-0199476P.
PR 26-APR-2000; 2000US-0199880P.
PR 26-APR-2000; 2000US-0200024P.
PR 26-APR-2000; 2000US-0200025P.
PR 09-JUN-2000; 2000US-0210809P.
PR 03-JUL-2000; 2000US-0215855P.
PR 17-JUL-2000; 2000US-0218591P.
PR 11-AUG-2000; 2000US-0224610P.
PR 27-FEB-2001; 2001US-0271814P.
XX (GERL/) GERLACH V L.
PA (ELLE/) ELLERMAN K.
PA (MACD/) MACDOUGALL J R.
PA (SMIT/) SMITHSON G.
XX Gerlach VL, Ellerman K, Macdougall JR, Smithson G;
PI WPI; 2003-066815/06.
XX N-PSDB; ABX56479.
XX Novel polypeptides and nucleic acids which are members of epidermal
PT growth factor, complement receptor families for diagnosing and treating
PT psychiatric conditions, depression, stroke, Alzheimer's and Parkinson's
PT disease.
XX Claim 1; Page 38; 91pp; English.
XX The invention describes an isolated POLYX (POLY1-17) polypeptide and its
CC variant. POLYX polypeptides (especially POLY5, POLY6 and POLY7), the
CC polynucleotides encoding them (I) and an anti-POLYX-antibody (III) are
CC useful for treating or preventing a pathology associated with POLYX
CC polypeptide in humans and for treating a syndrome associated with human
CC disease. POLYX polypeptide is also useful for identifying an agent that
CC binds to POLYX and a cell expressing POLYX is useful for identifying a
CC therapeutic agent for use in treatment of a pathology related to aberrant
CC expression or physiological interactions of the polypeptide. (III) is
CC useful for treating a pathological state in a mammal and for determining
CC the presence or amount of POLYX in a sample. POLY1-4 (GABA receptor-like
CC proteins) are useful for the treatment of psychiatric and medical
CC conditions, depression, stroke, Parkinson's disease, Huntington's
CC disease, Tourette's syndrome, amyotrophic lateral sclerosis, head trauma,
CC Alzheimer's disease, alcoholism, vigilance, anxiety, muscle tension,
CC epileptogenic activity and memory functions, cardiomyopathy and
CC arrhythmogenic right ventricular dysplasia. POLY5-8 (Epidermal growth
CC factor like proteins) may be useful for treating cancer, aberrant
CC angiogenesis, renal disease and diabetes. POLY12 (haematopoietic stem and
CC progenitor cell like protein) may be useful for treatment of leukaemia,
CC lupus and anaemia. POLY13 (sulfotransferase-like protein) may be useful
CC for treating cirrhosis, cholangitis, hepatitis, ulcers, hyperthyroidism
CC and developmental disorders. POLY14-16 (Syntactin-like proteins) may be
CC useful in treatment of Lambert-Eaton myasthenic syndrome, asthma, myxoid
CC liposarcoma and acute myeloid leukaemia, and POLY 18 may be useful in
CC treatment of cancers. Cells comprising (I) are useful for producing non-
CC human transgenic animals which are useful for studying the function
CC and/or activity of POLYX protein and for identifying and/or evaluating
CC modulators of POLYX protein activity. This is the amino acid sequence of
XX a novel human protein
SQ Sequence 294 AA;
Query Match 9.0%; Score 96; DB 6; Length 294;
Best Local Similarity 23.8%; Pred. No. 0.15;
Matches 44; Conservative 31; Mismatches 58; Indels 52; Gaps 10;
QY 27 BEINKAIDATAAEQSTID-PMKV-----PHADKFERHGVDF-----KGE- 70
DB 103 EYINSLNDLVKVKSEVNGSPSVTRILKSOHAAVF-RHFQOIMFYNDIAKQEK 161

QY 71 ---LAWNIEARG-----LKQMKRQGDANYKGE-----GIVKAHL---LIGVHDDIVS 113
Db 162 CKTFILRQLEVAGKEMSEEDVNDMLHQKWEVFNESLLTEINITKAQLSEIEQRKELVN 221
QY 114 MEYDLAYKLGDLHPTHTVTSIDQDFVALSLSEISDEG-----NITMTSFEVQFANVNH 169
Db 222 LE-----NQKDLRDLFIQISLLVEEGESINNIENTVNSTREYVNTTKEK 267
QY 170 GGLSI 174
Db 268 FGLAV 272
RESULT 15
ADB82715
ID ADB82715 standard; protein; 294 AA.
XX
AC ADB82715;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human protein sequence useful for the treatment of cancer (SeqID 1496).
XX human; prostate; cancer; cytostatic; gene therapy; vaccine;
XX immune response.
XX Homo sapiens.
XX WO2003050236-A2.
XX 19-JUN-2003.
XX
XX 04-SEP-2002; 2002WO-US028214.
XX
XX 07-DEC-2001; 2001US-00012697.
XX (CHIR) CHIRON CORP.
XX (HYSE-) HYSEQ INC.
XX Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI Crkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D;
PI Garcia V, Jones LW, Stache-Grain B, Scott EM;
XX WPI; 2003-513972/48.
XX New polynucleotides derived from human prostate, useful for modulating
PT immune response to prevent or treat cancer.
XX
XX Claim 18; SEQ ID NO 1496; 188pp; English.
XX This invention relates to novel isolated polynucleotides of human origin,
CC particularly isolated from the human prostate. Specifically, it refers to
CC the diagnostics and therapeutics comprising these novel human
CC polynucleotides, and includes the derived probes, antisense
CC oligonucleotides and antibodies thereof. The identification of these
CC human prostate genes that can inhibit tumour growth is useful for
CC understanding the progression and nature of complex diseases such as
CC cancer, and hence they are important in the drug discovery process. The
CC present invention describes these polynucleotides and encoded
CC polypeptides as exhibiting cytostatic activity, and through gene therapy
CC and/or vaccines they can be used to modulate the immune response for the
CC prevention or treatment of cancers, particularly of the prostate, but
CC also for breast, lung and colon cancer. This polypeptide sequence is a
CC human protein sequence useful for the treatment of cancer, used in an
CC exemplification of the invention. NOTE: These sequences are not given in
CC the specification but are provided on the WIPO website.
XX
SQ Sequence 294 AA;
Query Match 9.0%; Score 96; DB 7; Length 294;
Best Local Similarity 23.8%; Pred. No. 0.15;
Matches 44; Conservative 31; Mismatches 58; Indels 52; Gaps 10;

QY 27 EEINKAIDDAIAAEQSETID-PMKV-----PDHAKFERHVGIVDF-----KGE- 70
 Db 103 EYNRSNDLVKEVKSEVNGSSVTRILKSHQAMF-RHFQOIMFYNDTIAAKQEK 161
 QY 71 ---LAMENIEARG-----LKQMKRQGDANVKGE-----GIVKAHL--LIGVHDDIVS 113
 Db 162 CKTFILRQLEVAGKEMSEEDVNDMLHQKWEVFNESLLTEINITKAQLSEIQRHKELVN 221
 QY 114 MEYDLAYKGLDHPHTTHVISIDQFVVALSLEISDEG-----NITMTSEFEVQFANVNH 169
 Db 222 LE-----NQIKDLRLDFIQISLLVEQGESINNIENTVNSTKEYVNTKEK 267
 QY 170 GGLSI 174
 Db 268 FGLAV 272

RESULT 16
 AAE38817
 ID AAE38817 standard; protein; 294 AA.
 XX
 AC AAE38817;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human syntaxin-like protein (POLY16).
 XX
 KW Human; tumour; inflammatory disorder; vaccine; gene therapy; cytostatic;
 KW syntaxin-like protein; POLY16.
 OS Homo sapiens.
 XX
 PN US2003050232-A1.
 XX
 PD 13-MAR-2003.
 XX
 PF 19-APR-2001; 2001US-00839446.
 XX
 PR 19-APR-2000; 2000US-0198293P.
 PR 20-APR-2000; 2000US-0198645P.
 PR 25-APR-2000; 2000US-0199476P.
 PR 26-APR-2000; 2000US-0199880P.
 PR 26-APR-2000; 2000US-0200024P.
 PR 26-APR-2000; 2000US-0200025P.
 PR 09-JUN-2000; 2000US-0210809P.
 PR 17-JUL-2000; 2000US-0218591P.
 PR 11-AUG-2000; 2000US-0224610P.
 PR 27-FEB-2001; 2001US-0271814P.
 XX
 PA (TAUP/) TAUPIER R J.
 PA (PADI/) PADIGARU M.
 PA (SPYT/) SPYTEK K A.
 PA (BURG/) BURGESS C E.
 PA (VERN/) VERNET C A M.
 PA (FERN/) FERNANDES E R.
 PA (SHIM/) SHIMKETS R A.
 PA (LIUX/) LIU X.
 PA (MAJU/) MAJUMDER K.
 PA (COLM/) COLMAN S D.
 PA (ZERH/) ZERHUSEN B D.
 XX
 PI Taupier RJ, Padigaru M, Spytek KA, Burgess CE, Vernet CAM;
 PI Fernandes ER, Shinkets RA, Liu X, Majumder K, Colman SD;
 PI Zerhusen BD;
 XX
 DR WPI; 2003-605764/57.
 DR N-PSDB; AAD58968.
 XX
 XX New POLYX nucleic acid, useful for preparing a composition for treating
 PT or preventing e.g., tumor or inflammatory disorder.
 XX
 PS Claim 1; Page 41; 75pp; English.

XX
 CC The invention relates to new POLYX nucleic acid useful for preparing a
 CC composition for treating or preventing tumour or inflammatory disorder.
 CC The invention is useful as vaccine and in gene therapy. The nucleic acid
 CC is useful for preparing a composition for treating or preventing e.g.,
 CC tumour or inflammatory disorder. The present sequence is human syntaxin-
 CC like protein (POLY16)
 XX
 SQ Sequence 294 AA;
 Query Match 9.0%; Score 96; DB 7; Length 294;
 Best Local Similarity 23.8%; Pred. NO. 0.15;
 Matches 44; Conservative 31; Mismatches 58; Indels 52; Gaps 10;
 QY 27 EEINKAIDDAIAAEQSETID-PMKV-----PDHAKFERHVGIVDF-----KGE- 70
 Db 103 EYNRSNDLVKEVKSEVNGSSVTRILKSHQAMF-RHFQOIMFYNDTIAAKQEK 161
 QY 71 ---LAMENIEARG-----LKQMKRQGDANVKGE-----GIVKAHL--LIGVHDDIVS 113
 Db 162 CKTFILRQLEVAGKEMSEEDVNDMLHQKWEVFNESLLTEINITKAQLSEIQRHKELVN 221
 QY 114 MEYDLAYKGLDHPHTTHVISIDQFVVALSLEISDEG-----NITMTSEFEVQFANVNH 169
 Db 222 LE-----NQIKDLRLDFIQISLLVEQGESINNIENTVNSTKEYVNTKEK 267
 QY 170 GGLSI 174
 Db 268 FGLAV 272

RESULT 17
 AAE38815
 ID AAE38815 standard; protein; 294 AA.
 XX
 AC AAE38815;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human syntaxin-like protein (POLY14).
 XX
 KW Human; tumour; inflammatory disorder; vaccine; gene therapy; cytostatic;
 KW syntaxin-like protein; POLY14; chromosome 1.
 OS Homo sapiens.
 XX
 PN US2003050232-A1.
 XX
 PD 13-MAR-2003.
 XX
 PF 19-APR-2001; 2001US-00839446.
 XX
 PR 19-APR-2000; 2000US-0198293P.
 PR 20-APR-2000; 2000US-0198645P.
 PR 25-APR-2000; 2000US-0199476P.
 PR 26-APR-2000; 2000US-0199880P.
 PR 26-APR-2000; 2000US-0200024P.
 PR 26-APR-2000; 2000US-0200025P.
 PR 09-JUN-2000; 2000US-0210809P.
 PR 17-JUL-2000; 2000US-0218591P.
 PR 11-AUG-2000; 2000US-0224610P.
 PR 27-FEB-2001; 2001US-0271814P.
 XX
 PA (TAUP/) TAUPIER R J.
 PA (PADI/) PADIGARU M.
 PA (SPYT/) SPYTEK K A.
 PA (BURG/) BURGESS C E.
 PA (VERN/) VERNET C A M.
 PA (FERN/) FERNANDES E R.
 PA (SHIM/) SHIMKETS R A.
 PA (LIUX/) LIU X.
 PA (MAJU/) MAJUMDER K.
 PA (COLM/) COLMAN S D.

(ZERH/) ZERHUSEN B D.

Taupier RJ, Padigar M, Spytek KA, Burgess CE, Vernet CAM;
 Fernandes ER, Shimkets RA, Liu X, Majumder K, Colman SD;
 Zerhusen BD;
 WPI; 2003-605764/57.
 N-PSDB; AAD58966.

New POLYX nucleic acid, useful for preparing a composition for treating
 or preventing e.g., tumor or inflammatory disorder.

Claim 1; Page 38; 75pp; English.

The invention relates to new POLYX nucleic acid useful for preparing a
 composition for treating or preventing tumour or inflammatory disorder.
 The invention is useful as vaccine and in gene therapy. The nucleic acid
 is useful for preparing a composition for treating or preventing e.g.,
 tumour or inflammatory disorder. The present sequence is human syntaxin-
 like protein (POLY14). POLY14 gene is located on chromosome 1. Note: The
 present sequence is encoded by a DNA containing translational exceptions

Sequence 294 AA;

Query Match 9.0%; Score 96; DB 7; Length 294;
 Best Local Similarity 23.8%; Pred. NO. 0.15;
 Matches 44; Conservative 31; Mismatches 58; Indels 52; Gaps 10;

27 BEINKAIDATAAIBQSTID-PKKV-----PHADKFEHVGIVDF-----KGE- 70
 103 EYINSLNDLVKVEKSEVNGSPSVTRILKSHRAAMP-RHQOIMFYNDTTAAKQEK 161
 71 ---LAMRNIEARG-----LKMQRQGDANVKGEE-----GIYKAHL--LIGVHDDIVS 113
 162 CKTFILRQLEVAGKEMSEEDVNDMLNQKWEVFNESLILTEINIKAQLSIEIQSHKELVN 221
 114 MEYDLAYKLGDLHPHTHVISIQDFVALSLEISDEG-----NITMTSEVRFQFANVNH 169
 222 LE-----NQIKRLDLFIQISLLVBEQGESINNIENTVNSTKEYVNTNKEK 267

170 GGLSI 174
 268 FGLAV 272

RESULT 18
 AAU19614
 ID AAU19614 standard; protein; 314 AA.

AC AAU19614;
 DT 04-DEC-2001 (first entry)
 XX Human diagnostic and therapeutic polypeptide (DITHP) #200.
 DE Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;
 KW cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;
 KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;
 KW respiratory disorder.
 XX Homo sapiens.
 OS WO200162927-A2.
 PN 30-AUG-2001.
 PD 21-FEB-2001; 2001WO-US006059.
 XX 24-FEB-2000; 2000US-0184693P.
 PR 24-FEB-2000; 2000US-0184697P.
 PR 24-FEB-2000; 2000US-0184698P.
 PR 24-FEB-2000; 2000US-0184768P.
 PR 24-FEB-2000; 2000US-0184769P.

24-FEB-2000; 2000US-0184770P.
 24-FEB-2000; 2000US-0184771P.
 24-FEB-2000; 2000US-0184772P.
 24-FEB-2000; 2000US-0184773P.
 24-FEB-2000; 2000US-0184774P.
 24-FEB-2000; 2000US-0184776P.
 24-FEB-2000; 2000US-0184777P.
 24-FEB-2000; 2000US-0184797P.
 24-FEB-2000; 2000US-0184813P.
 24-FEB-2000; 2000US-0184837P.
 24-FEB-2000; 2000US-0184841P.
 24-FEB-2000; 2000US-0185213P.
 24-FEB-2000; 2000US-0185216P.
 12-MAY-2000; 2000US-0203785P.
 15-MAY-2000; 2000US-0204226P.
 16-MAY-2000; 2000US-0204525P.
 16-MAY-2000; 2000US-0204821P.
 16-MAY-2000; 2000US-0204908P.
 16-MAY-2000; 2000US-0205232P.
 17-MAY-2000; 2000US-0204815P.
 17-MAY-2000; 2000US-0204863P.
 17-MAY-2000; 2000US-0205221P.
 17-MAY-2000; 2000US-0205285P.
 17-MAY-2000; 2000US-0205288P.
 17-MAY-2000; 2000US-0205287P.
 17-MAY-2000; 2000US-0205323P.
 17-MAY-2000; 2000US-0205324P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;
 PI Chen A, D'sa SA, Amshay S, Dahl CR, Dam TC, Daniels SE, Dufour GE;
 PI Flores V, Fong WT, Greenawalt LB, Hillman JL, Jones AL, Liu TF;
 PI Roseberry AM, Rosen BH, Russo FD, Stockdreher TK, Daifo A;
 PI Wright RJ, Yap BP, Yu JY, Bradley DL, Bratcher SR, Chen W;
 PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;
 XX WPI; 2001-502867/55.
 DR N-PSDB; AAS31185.
 XX Polynucleotides encoding diagnostic and therapeutic proteins, e.g.
 PT enzymes, hormones and receptors, useful in diagnostics and therapeutics.
 PT Claim 27; Page 516; 522pp; English.
 XX The invention relates to polynucleotides (I) encoding diagnostic and
 CC therapeutic (DITHP) polypeptides (II), which include e.g. enzymes, and
 CC proteins involved in growth and development and receptors. (I) and (II)
 CC may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate DITHP expression. For example, (I) and (II)
 CC may be used to treat disorders associated with decreased polypeptide
 CC expression by rectifying mutations or deletions in a patient's genome,
 CC that affect the activity of the DITHPs, by expressing inactive proteins
 CC or supplementing the patient's own production of them. (I) and (II) may
 CC be used to treat diseases, for example, cell proliferative disorder,
 CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,
 CC leukaemia, autoimmune disorders, and respiratory disorders. Additionally,
 CC (I) may be used to produce the DITHPs, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the protein. (I) and
 CC its complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acids in
 CC samples, and therefore which patients may be in need of restorative
 CC therapy. (II) may also be used as antigens in the production of
 CC antibodies against DITHPs and in assays to identify modulators of DITHP
 CC expression and activity. The anti-DITHP antibodies and antagonists may
 CC also be used to down regulate expression and activity. The anti-DITHP
 CC antibodies may also be used as diagnostic agents for detecting the
 CC presence of DITHPs in samples (e.g. by enzyme linked immunosorbent assay
 CC (ELISA)). AAU19415-AAU19625 represent human diagnostic and therapeutic
 CC (DITHP) polypeptides of the invention
 XX Sequence 314 AA;
 SQ

CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
XX SQ Sequence 270 AA;

Query Match 8.6%; Score 91.5; DB 6; Length 270;
Best Local Similarity 25.5%; Pred.No.0.4;
Matches 47; Conservative 30; Mismatches 70; Indels 37; Gaps 11.

QY 26 TEEINKATDDAIAIEQSE--TIDPMKVPHAD-----KFERHVGIVDFKG---ELAVEN 75
Db 26 SEELGTGDECLHDHDESEIGAFDNR--DGDRRDTHFAVLGLSDVSGNINETATRI 83
QY 76 TEAR-----GLKQMQRGGDANYKGEIGVKAHLILGVHDDIVSMEDYLAYKLGDLPHT 129
Db 84 VEERERESFASLDDLARRVD--LSGEE--VEALALAGAPDLLVGSRRGALWQIQINGVA 139
QY 130 HWISDIQDFVA-----LSLEISDEG----NITMTSFEVRQFANVVNHIGLSILD 177
Db 140 PQQLDVQ--VVTQPPLPEPTQMELLGDLRATGISTADHPVRQVRDALNRRGVQGV-DR 196
QY 178 IFGV 181
Db 197 LDGV 200

RESULT 23
ID ARM64952
ARM64952 standard; protein; 1148 AA.
AC ARM64952;
XX DT 20-OCT-2003 (first entry)
DE Propionibacterium acnes immunogenic polypeptide #29628.
XX DE
XX KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine, immunogenic.
XX CS Propionibacterium acnes.
XX OS
XX PN WO2003033515-A1.
XX PD
XX PP 24-APR-2003.
XX PF 11-OCT-2002; 2002WO-US032727.
XX PR 15-OCT-2001; 2001US-00978825.
XX PA {CORI-} CORIXA CORP.
XX PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Valliee-Douglais J;
XX WPI; 2003-381789/36.
XX DR
XX PT New propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX PS Claim 7; SEQ ID NO 29628; 1481bp; English.
XX CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)

QY 196 KEMTKVLAPA 205
Db 1334 --TRVLTDA 1340

RESULT 25
AAE38816
ID AAE38816 standard; protein; 294 AA.
XX AC AAE38816;
XX DT 18-DEC-2003 (first entry)
XX DE Human syntaxin-like protein (POLY15).
XX KW Human; tumour; inflammatory disorder; vaccine; gene therapy; cytostatic;
XX KW syntaxin-like protein; POLY15.
XX OS Homo sapiens.
XX PN US2003050232-A1.
XX PD 13-MAR-2003.
XX PF 19-APR-2001; 2001US-00839446.
XX PR 19-APR-2000; 2000US-0198293P.
XX PR 20-APR-2000; 2000US-0198645P.
XX PR 25-APR-2000; 2000US-0199476P.
XX PR 26-APR-2000; 2000US-0199880P.
XX PR 26-APR-2000; 2000US-0200024P.
XX PR 26-APR-2000; 2000US-0200025P.
XX PR 09-JUN-2000; 2000US-0210809P.
XX PR 17-JUL-2000; 2000US-0218591P.
XX PR 11-AUG-2000; 2000US-0224610P.
XX PR 27-FEB-2001; 2001US-0271814P.

PA (TAUP/) TAUPIER R J.
PA (PADI/) PADIGARU M.
PA (SPYT/) SPYTEK K A.
PA (BURG/) BURGESS C E.
PA (VERN/) VERNET C A M.
PA (FERN/) FERNANDES E R.
PA (SHIM/) SHINKETS R A.
PA (LIUX/) LIU X.
PA (MAJU/) MAJUMDER K.
PA (COLM/) COLMAN S D.
PA (ZERH/) ZERHUSEN B D.

XX Taupier RJ, Padigaru M, Spytek KA, Burgess CE, Vernet CAM;
PI Fernandes ER, Shinkets RA, Liu X, Majumder K, Colman SD;
PI Zerhusen BD;
XX WPI; 2003-605764/57.
XX DR N-PSDB; AAD58967.
XX PT New POLYX nucleic acid, useful for preparing a composition for treating
XX or preventing e.g., tumor or inflammatory disorder.
XX Claim 1; Page 40; 75pp; English.
XX CC The invention relates to new POLYX nucleic acid useful for preparing a
XX composition for treating or preventing tumour or inflammatory disorder.
XX CC The invention is useful as vaccine and in gene therapy. The nucleic acid
XX is useful for preparing a composition for treating or preventing e.g.,
XX tumour or inflammatory disorder. The present sequence is human syntaxin-
XX like protein (POLY15)
XX SQ Sequence 294 AA;

Query Match 8.5%; Score 91; DB 7; Length 294;
Best Local Similarity 23.5%; Pred. No. 0.51;
Matches 44; Conservative 55; Indels 56; Gaps 11;

QY 27 BEINKAIDATAAIEQSETID-PMKV-----POHADKFERHVGIVDF-----KGE- 70
Db 103 EYINRSLNDLVEKVKSEVNGPSVVTIRIKSQHAF-RHQOIMFIYNDTIAAKQEK 161
QY 71 ---LAMRNIEARGLKQMKRQGDANYKBEG-----IVKAHL--LIGVHDDI 111
Db 162 CKTFILRQLEVAG-KEMSEE-DVNDNLHQKWEVFNESLLTEINITKAQLSEIQRHKEK 219
QY 112 VSMEDYDLAYKGLDLPHTTHVJSDIQDFVVALSLEISDEG---NITWTSFEVRQAPVNVN 167
Db 220 VNLE-----NQIKDLRDLFIQTSLLVEEQGESINNIENTVNSTKEYVNTK 265
QY 168 HIGGLSI 174
Db 266 EKFLAV 272

RESULT 26
ABU19882
ID ABU19882 standard; protein; 944 AA.
XX AC ABU19882;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #5409.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Borrelia cepacia.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX DR N-PSDB; ACA23752.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 47806; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 944 AA;

Query Match 8.5%; Score 90.5; DB 6; Length 944;
 Best Local Similarity 27.1%; Pred. No. 3;
 Matches 29; Conservative 19; Mismatches 52; Indels 7; Gaps 2;
 QY 45 TIDPMKVPDHADKFRHHVGIYDFK----GELAMNIEARGLKQMKROGDANVKEEGIVK 100
 DB 357 TLDIVGAPVNGANANRPIRTIRFERDRMGNLKVQHTPTFTRFERDKGDRTVKVRPTTP 416
 QY 101 AHLLGVHDDIVSMEDYDLAYKLGDLH---PTTHVISIDQDFVALSL 144
 DB 417 AGIALGIPDVVSFFYDXAGRLMAEHGVNGIVEYLDALDNVTTLAL 463
 RESULT 27
 ID AAB18324
 XX AAB18324 standard; protein; 1558 AA.
 AC AAB18324;
 XX
 DT 07-NOV-2000 (first entry)
 XX
 DE Plasmodium falciparum chromosome 2 related protein SEQ ID NO:182.
 XX
 KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
 KW antimalarial; malaria; protozoacide; infection; insecticide.
 XX
 OS Plasmodium falciparum.
 XX
 PN WO200025728-A2.
 XX
 PD 11-MAY-2000.
 XX
 PF 05-NOV-1999; 99WO-US026796.
 PR 05-NOV-1999; 98US-0107131P.
 XX
 PA (HOFF/) HOFFMAN S.
 PA (CARU/) CARUCCI D.
 PA (GARD/) GARDNER M.
 PA (VENT/) VENTER J C.
 XX
 PI Hoffman S, Carucci D, Gardner M, Venter JC;
 XX
 DR WPI; 2000-365347/31.
 XX
 PT Proteins encoded by chromosome 2 of the human malarial parasite,
 PT Plasmodium falciparum, useful as antimalarial vaccines and in the
 PT diagnosis of *P. falciparum* infection.
 XX
 PS Disclosure; Page 410-414; 577pp; English.
 XX
 CC The present invention describes proteins and their fragments (I) encoded
 CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
 CC Also described are: (1) nucleotide sequences (II) encoding (I); and (2)

CC vaccines against *P. falciparum* infection comprising (I) or (II). (I) and
 CC (II) are useful for the development of vaccines against *P. falciparum*
 CC infection. (I) and polyclonal antisera or a monoclonal antibody raised to
 CC immunogens comprising the sequences of (I), are useful in the detection
 CC of infection with *P. falciparum*. Furthermore, (I) (especially when they
 CC are rifins or secreted or membrane proteins) can aid the identification
 CC of drugs to treat or prevent *P. falciparum* infection, or they can be used
 CC to identify drug resistance in *P. falciparum*. Sequencing of the
 CC Plasmodium chromosome 2 and the subsequent identification of proteins
 CC encoded by it will help to expand our understanding of parasite biology,
 CC a process hampered by the complexity of the parasitic lifecycle, and
 CC provide new targets for vaccine and drug development. Parasite resistance
 CC to drugs and mosquito resistance to insecticides have led to a resurgence
 CC of malaria in many parts of the world, and there is a pressing need for
 CC vaccines and new drugs. AAB70078 to AAB70287 and AAB18144 to AAB18352
 CC represent nucleotide and protein sequences given in the present
 CC invention, but which are not specifically mentioned within the
 CC specification
 XX
 XX Sequence 1558 AA;

Query Match 8.4%; Score 90; DB 3; Length 1558;
 Best Local Similarity 23.3%; Pred. No. 7;
 Matches 47; Conservative 43; Mismatches 78; Indels 34; Gaps 9;
 QY 22 YDKITEINKALDDAIAIEQSETIDPMKVPDHADKFRHHVGIYDFKELAMNIEARGL 81
 DB 747 FNTVLDKVEETVEISGESLENNE---MDKAFFSEIFDNVKGIOENLTGMFRSLETIV 802
 QY 82 KQMKROGDANVKEEGIVKAHLIGVHDDIVSMEDYDLAYKLGDLHPTTHVISIDQDFVVA 141
 DB 803 IQSEKVDLN---ENVVSSIL-----DNIENKGLNKLKLENISSTEGVQETVTEHV-- 851
 QY 142 LSLEISDEGNITWTSFEV---RQFANVNHIGLS---ILDPIFGVLSVLTG--IFQ 191
 DB 852 -----EQNV-YVDVDPAMKQDFLGILNEAGLKEMFNLEDVFKESDVITVEEIKD 903
 QY 192 DTVRKEMTKVLAPAKRELEKN 213
 DB 904 EPVQKEVEKETVSIIE-EMEN 924

RESULT 28
 ID AAW24790
 XX AAW24790 standard; protein; 1786 AA.
 AC AAW24790;
 XX
 DT 08-OCT-1997 (first entry)
 XX
 DE *P. falciparum* liver stage antigen-3.
 XX
 KW Plasmodium falciparum; pre-erythrocyte; liver stage antigen; serum;
 KW prophylaxis; thai strain; gene organisation; exon; intron; hydrophobic;
 KW glycosyl-phosphatidylinositol membrane anchoring sequence; antibody;
 XX vaccine; immunotherapy; malaria.
 OS Plasmodium falciparum.
 XX
 PH Key Location/Qualifiers
 FT Region 223..278
 FT /note="repeat region 1"
 FT Region 279..818
 FT /note="repeat region 2"
 FT Region 1537..1576
 FT /note="repeat region 3"
 XX
 PN WO9641877-A2.
 XX
 XX 27-DEC-1996.
 XX
 PF 12-JUN-1996; 96WO-FR000894.

QY	6	LIAAVAFVWSAD-----PIH--YDKITEIN-----KAIDATAAIEQS	43
	:	: : : :	:
DB	217	VLAIPAFILINSNVIGLNLPPAMFPTTHIRYSALLEMKQHTKTFGSVKAIENVCLRLNAG	276
	:	: : : :	:
QY	44	ETID-----PMKVDPHADKFERHVGVDFKGELAMRNTFARGLKQWKMGQDANV	92
	:	: : : :	:
DB	277	EIVSLGCGSGSKSTLMKVLGGIYPHGSYEGEIIFAGE---EIQASHIRDTERKGIALL	332
	:	: : : :	:
QY	93	KGESGIVK-----AHLJLG----VHDDIVSMBYDLAYLGLDLPHTPHVTSIDTQDFVVALS	144
	:	: : : :	:
DB	333	HQEALVKELTVLENIPLNGEITHNGI--MDYDL-----MTLRQCKLLAQVSL	378
	:	: : : :	:
QY	145	EISDE---GNITMTSFEVRQANVNHHIGLSILDPIFGVLSLDVLTATFQDTVR	195
	:	: : : :	:
DB	379	SISPTRVGDLGQQQLVEIKALKNQVRLIIDEPASUTEGETSILLDIIR	432
	:	: : : :	:
RESULT 31			
AAB58737	ID	AAB58737 standard; protein; 405 AA.	
XX	AC	AAB58737;	
XX	DT	27-MAR-2001 (first entry)	
XX	XX	Breast and ovarian cancer associated antigen protein sequence SEQ ID 445.	
XX	DE	Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;	
KW	KW	nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;	
KW	KW	antidiabetic; antiinflammatory; antitumor; vulnary; anticonvulsant;	
KW	KW	antibacterial; antifungal; antiparasitic; cardiant; immune disorder;	
KW	KW	Addison's disease; allergy; autoimmune haemolytic anaemia;	
KW	KW	autoimmune thyroiditis; diabetes mellitus; Crohn's disease;	
KW	KW	multiple sclerosis; rheumatoid arthritis; ulcerative colitis;	
KW	KW	cardiovascular disorder; wound healing; neurological disease.	
OS	OS	Homo sapiens.	
XX	XX	WO2000055173-A1.	
PN	PN	21-SEP-2000.	
XX	PD	08-MAR-2000; 2000MO-US005881.	
PF	PF	12-MAR-1999; 99US-0124270P.	
XX	PR	(HUMA-) HUMAN GENOME SCI INC.	
PA	PA	Rosen CA, Ruben SM;	
XX	PI	WPI; 2000-611515/58.	
XX	PP	N-PSDB; AAF21640.	
DR	DR	New human breast and ovarian cancer associated gene sequences and the	
DR	DR	polypeptides encoded by these genes, useful in the prevention, treatment	
XX	PT	and diagnosis of cancer, immune disorders, cardiovascular disorders and	
XX	PT	neurological diseases.	
XX	PS	Claim 11; Page 872-874; 1299pp; English.	
CC	CC	Sequences AAF21614 - AAF22031 represent DNA sequences encoding human	
CC	CC	proteins AAB58711 - AAB59128. The DNA and protein sequences are	
CC	CC	associated with breast and ovarian cancer. Included in the invention are	
CC	CC	sequences AAF22032 - AAF22040 and AAB59129 which are used in the	
CC	CC	isolation and characterisation of the DNA and protein sequences of the	
CC	CC	invention. The breast and ovarian cancer associated DNA, protein, agonist	
CC	CC	or antagonist sequences exhibit cytostatic; immunosuppressive; nootropic;	
CC	CC	neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;	
CC	CC	antiinflammatory; antitumor; vulnary; anticonvulsant; antibacterial;	
CC	CC	antifungal; antiparasitic and cardiant activity. The polynucleotide and	
CC	CC	protein sequences are used in the diagnosis of cancer, particularly	
CC	CC	breast and ovarian cancer. The nucleic acid sequences, proteins, agonists	
CC	CC	and agonists may also be used in the diagnosis, prevention and treatment	

CC of immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC cardiovascular disorders such as myocardial ischaemia; wound healing;
CC neurological diseases such as cerebral anoxia and epilepsy; and
CC infectious diseases
XX
SQ Sequence 405 AA;

Query Match 8.3%; Score 88.5; DB 3; Length 405;
Best Local Similarity 24.6%; Pred. No. 1.5;
Matches 30; Conservative 25; Mismatches 50; Indels 17; Gaps 3;

QY 18 DPHYDKITEINKAIDDAIAAEQSETI-----DPMKVPDHD-KFERHV 62
Db 21 DKFSFDLGKGEVIAKWDIAIAATMKVGEVCHITCKPEYAGSAGSPKPPNATLVFE--V 78
QY 63 GIVDFKGLAMRNIEARGLQKMQGDANVKGEIGVKAHLIGVHDDIVSMEYDLAYKL 122
Db 79 ELFEFKGEDLTREEDGGIIRRIQTRGEGYAKPNEGAIVEVALEGYVYKDKLPQRELRFEI 138
QY 123 GD 124
Db 139 GE 140

RESULT 32
AD60975
ID ADE60975 standard; protein; 458 AA.
XX AC ADE60975;
XX DT 29-JAN-2004 (first entry)
XX DE Human Protein Q02790, SEQ ID NO 6889.
XX KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX OS Homo sapiens.
XX FN WO2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX FI Woolf C, D'urso D, Befort K, Costigan M;
XX DR WPI; 2003-268312/26.
XX DR GENBANK; Q02790.
XX PT New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX FS Claim 1; Page; 1017pp; English.
XX
The invention discloses a composition comprising two or more isolated rat
or human polynucleotides or a polynucleotide which represents a fragment,
derivative or allelic variation of the nucleic acid sequence. Also
claimed are a vector comprising the novel polynucleotide, a host cell
comprising the vector, a method for identifying a nucleotide sequence
which is differentially regulated in an animal subjected to pain and a
kit to perform the method, an array, a method for identifying an agent
that increases or decreases the expression of the polynucleotide sequence

CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. the polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 458 AA;

Query Match 8.3%; Score 88.5; DB 7; Length 458;
Best Local Similarity 24.6%; Pred. No. 1.8;
Matches 30; Conservative 25; Mismatches 50; Indels 17; Gaps 3;

QY 18 DPHYDKITEINKAIDDAIAAEQSETI-----DPMKVPDHD-KFERHV 62
Db 74 DKFSFDLGKGEVIAKWDIAIAATMKVGEVCHITCKPEYAGSAGSPKPPNATLVFE--V 131
QY 63 GIVDFKGLAMRNIEARGLQKMQGDANVKGEIGVKAHLIGVHDDIVSMEYDLAYKL 122
Db 132 ELFEFKGEDLTREEDGGIIRRIQTRGEGYAKPNEGAIVEVALEGYVYKDKLPQRELRFEI 191
QY 123 GD 124
Db 192 GE 193

RESULT 33
ADE60960
ID ADE60960 standard; protein; 458 AA.
XX AC ADE60960;
XX DT 29-JAN-2004 (first entry)
XX DE Human Protein Q02790, SEQ ID NO 6874.
XX KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX OS Homo sapiens.
XX FN WO2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX FI Woolf C, D'urso D, Befort K, Costigan M;
XX DR WPI; 2003-268312/26.
XX DR GENBANK; Q02790.

PT New composition comprising two or more isolated polypeptides, useful for
 XX preparing a medicament for treating pain in an animal.
 PS Claim 1; Page; 1017pp; English.
 XX
 CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 458 AA;
 Query Match 8.3%; Score 88.5; DB 7; Length 458;
 Best Local Similarity 24.6%; Pred. No. 1.8;
 Matches 30; Conservative 25; Mismatches 50; Indels 17; Gaps 3;
 QY 18 DPIHYDKITEINKAIDDAIAAEQSETI-----DPMKVPHAD-KPERHV 62
 Db 74 DKFSDLGKGEVIAKWDIAIATMKVGEVCHITCKPEYAGSAGSPKIPPNATLVFE--V 131
 QY 63 GIVDFKGEIAMRNIEARGLKQKQGDANVKGEGIVKAHLIGVHDDIVSMYDLYAKL 122
 Db 132 ELFEFKGEDLTEEDGGIIRRIQTRGEGYAKPNEGAIVEALEGYKKDKLFDQRELARFEI 191
 QY 123 GD 124
 Db 192 GE 193
 RESULT 34
 ID ADE60963
 XX ADE60963 standard; protein; 458 AA.
 AC ADE60963;
 XX
 XX 29-JAN-2004 (first entry)
 DE Human Protein Q02790, SEQ ID NO 6877.
 XX
 KW Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX
 OS Homo sapiens.
 XX
 XX W02003016475-A2.
 XX
 XX 27-FEB-2003.
 XX
 XX 14-AUG-2002; 2002WO-US025765.
 XX

XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 PI Woolf C, D'urso D, Befort K, Costigan M;
 XX
 DR WPI; 2003-268312/26.
 DR GENBANK; Q02790.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX
 PS Claim 1; Page; 1017pp; English.
 XX
 CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 458 AA;
 Query Match 8.3%; Score 88.5; DB 7; Length 458;
 Best Local Similarity 24.6%; Pred. No. 1.8;
 Matches 30; Conservative 25; Mismatches 50; Indels 17; Gaps 3;
 QY 18 DPIHYDKITEINKAIDDAIAAEQSETI-----DPMKVPHAD-KPERHV 62
 Db 74 DKFSDLGKGEVIAKWDIAIATMKVGEVCHITCKPEYAGSAGSPKIPPNATLVFE--V 131
 QY 63 GIVDFKGEIAMRNIEARGLKQKQGDANVKGEGIVKAHLIGVHDDIVSMYDLYAKL 122
 Db 132 ELFEFKGEDLTEEDGGIIRRIQTRGEGYAKPNEGAIVEALEGYKKDKLFDQRELARFEI 191
 QY 123 GD 124
 Db 192 GE 193
 RESULT 35
 ID ADE60972
 XX ADE60972 standard; protein; 458 AA.
 AC ADE60972;
 XX
 XX 29-JAN-2004 (first entry)
 DE
 XX
 KW Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX
 OS Homo sapiens.
 XX
 XX W02003016475-A2.
 XX
 XX 27-FEB-2003.
 XX
 XX 14-AUG-2002; 2002WO-US025765.
 XX

Human Protein Q02790, SEQ ID NO 6886.
Human; pain; neuronal tissue; gene therapy;
spinal segmental nerve injury; chronic constriction injury; CCI;
spared nerve injury; SNI; Chung.
Homo sapiens.
WO2003016475-A2.
27-FEB-2003.
14-AUG-2002; 2002WO-US025765.
14-AUG-2001; 2001US-0312147P.
01-NOV-2001; 2001US-0346382P.
26-NOV-2001; 2001US-0333347P.
(GEO) GEN HOSPITAL CORP.
(FARB) BAYER AG.
Woolf C, D'urso D, Befort K, Costigan M;
WPI; 2003-268312/26.
GENBANK; Q02790.
New composition comprising two or more isolated polypeptides, useful for
preparing a medicament for treating pain in an animal.
Claim 1; Page; 1017pp; English.
The invention discloses a composition comprising two or more isolated rat
or human polynucleotides or a polynucleotide which represents a fragment,
derivative or allelic variation of the nucleic acid sequence. Also
claimed are a vector comprising the novel polynucleotide, a host cell
comprising the vector, a method for identifying a nucleotide sequence
which is differentially regulated in an animal subjected to pain and a
kit to perform the method, an array, a method for identifying an agent
that increases or decreases the expression of the polynucleotide sequence
that is differentially expressed in neuronal tissue of a first animal
subjected to pain, a method for identifying a compound which regulates
the expression of a polynucleotide sequence which is differentially
expressed in an animal subjected to pain, a method for identifying a
compound that regulates the activity of one or more of the
polynucleotides, a method for producing a pharmaceutical composition, a
method for identifying a compound or small molecule that regulates the
activity in an animal of one or more of the polypeptides given in the
specification, a method for identifying a compound useful in treating
pain and a pharmaceutical composition comprising the one or more
polypeptides or their antibodies. The polynucleotide or the compound that
modulates its activity is useful for preparing a medicament for treating
pain (e.g. spinal segmental nerve injury (SNI)), chronic constriction
injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene
therapy). The sequence presented is a human protein (shown in Table 2 of
the specification) which is differentially expressed during pain. Note:
The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic form directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences.
Sequence 458 AA;

Human Protein Q02790, SEQ ID NO 6886.
Human; pain; neuronal tissue; gene therapy;
spinal segmental nerve injury; chronic constriction injury; CCI;
spared nerve injury; SNI; Chung.
Homo sapiens.
WO2003016475-A2.
27-FEB-2003.
14-AUG-2002; 2002WO-US025765.
14-AUG-2001; 2001US-0312147P.
01-NOV-2001; 2001US-0346382P.
26-NOV-2001; 2001US-0333347P.
(GEO) GEN HOSPITAL CORP.
(FARB) BAYER AG.
Woolf C, D'urso D, Befort K, Costigan M;
WPI; 2003-268312/26.
GENBANK; Q02790.
New composition comprising two or more isolated polypeptides, useful for
preparing a medicament for treating pain in an animal.
Claim 1; Page; 1017pp; English.
The invention discloses a composition comprising two or more isolated rat
or human polynucleotides or a polynucleotide which represents a fragment,
derivative or allelic variation of the nucleic acid sequence. Also
claimed are a vector comprising the novel polynucleotide, a host cell
comprising the vector, a method for identifying a nucleotide sequence
which is differentially regulated in an animal subjected to pain and a
kit to perform the method, an array, a method for identifying an agent
that increases or decreases the expression of the polynucleotide sequence
that is differentially expressed in neuronal tissue of a first animal
subjected to pain, a method for identifying a compound which regulates
the expression of a polynucleotide sequence which is differentially
expressed in an animal subjected to pain, a method for identifying a
compound that regulates the activity of one or more of the
polynucleotides, a method for producing a pharmaceutical composition, a
method for identifying a compound or small molecule that regulates the
activity in an animal of one or more of the polypeptides given in the
specification, a method for identifying a compound useful in treating
pain and a pharmaceutical composition comprising the one or more
polypeptides or their antibodies. The polynucleotide or the compound that
modulates its activity is useful for preparing a medicament for treating
pain (e.g. spinal segmental nerve injury (SNI)), chronic constriction
injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene
therapy). The sequence presented is a human protein (shown in Table 2 of
the specification) which is differentially expressed during pain. Note:
The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic form directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences.
Sequence 458 AA;
Query Match 8.3%; Score 88.5; DB 7; Length 458;
Best Local Similarity 24.6%; Pred. NO. 1.8;
Matches 30; Conservative 25; Mismatches 50; Indels 17; Gaps 3;
QY 18 DPHVYDKTEINKAIDDAIAEQSETI-----DPMKVPDHAD-KFERHV 62
DB 74 DKFSDLGKGEVIAKWDIAIAIMKVECHITCKPEYAGSAGSPKIPNATLVE--V 131
QY 63 GIVDFKGLAMENIEARGKQKMGQDANVKGEGIVAHLLIGVHDIVSMVDLAYKL 122
DB 132 ELFEFKGDLTEEDGGIIRIOTRGEYAKPNEGAIIVEALEGYKDKLFDQELRFEI 191

Query Match 8.3%; Score 88.5; DB 7; Length 458;
 Best Local Similarity 24.6%; Pred. No. 1.8;
 Matches 30; Conservative 25; Mismatches 50; Indels 17; Gaps 3;

QY 18 DPHYDKITEEINKAIDDAIAIAESETI-----DPMKVPDHAD-KFERHV 62
 Db 74 DKFSFDLKGGEVKAWDIAIATMKVGEVCHITCKPEYAYGAGSPKIPPNATLVFE--V 131
 QY 63 GIVDFKGLAMENIARGLKQMKROGDANVKGEIGVKAHLLIGVHDDIVSMEYDLAYKL 122
 Db 132 ELFEFKGEDLTBEEDGGIIRRIQTRGEGYAKPNEGAIIVEVALEGGYKDKLFDQRELRFEI 191

QY 123 GD 124
 Db 192 GE 193

RESULT 37
 ADE60969
 ID ADE60969 standard; protein; 458 AA.
 XX AC ADE60969;
 XX DT 29-JAN-2004 (first entry)
 XX DE Human Protein Q02790, SEQ ID NO 6883.
 KW Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX OS Homo sapiens.
 XX PN WO2003016475-A2.
 XX PD 27-FEB-2003.
 XX PF 14-AUG-2002; 2002WO-US025765.
 XX PR 14-AUG-2001; 2001US-0312147P.
 XX PR 01-NOV-2001; 2001US-0346382P.
 XX PR 26-NOV-2001; 2001US-0333347P.
 XX PA (GEO) GEN HOSPITAL CORP.
 XX PA (FARB) BAYER AG.
 XX PI Woolf C, D'urso D, Befort K, Costigan M;
 XX WPI; 2003-268312/26.
 XX DR GENBANK; Q02790.

New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the

specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ffp.wipo.int/pub/published_pct_sequences.

XX Sequence 458 AA;

Query Match 8.3%; Score 88.5; DB 7; Length 458;
 Best Local Similarity 24.6%; Pred. No. 1.8;
 Matches 30; Conservative 25; Mismatches 50; Indels 17; Gaps 3;

QY 18 DPHYDKITEEINKAIDDAIAIAESETI-----DPMKVPDHAD-KFERHV 62
 Db 74 DKFSFDLKGGEVKAWDIAIATMKVGEVCHITCKPEYAYGAGSPKIPPNATLVFE--V 131
 QY 63 GIVDFKGLAMENIARGLKQMKROGDANVKGEIGVKAHLLIGVHDDIVSMEYDLAYKL 122
 Db 132 ELFEFKGEDLTBEEDGGIIRRIQTRGEGYAKPNEGAIIVEVALEGGYKDKLFDQRELRFEI 191
 QY 123 GD 124
 Db 192 GE 193

RESULT 38
 AAW54038
 ID AAW54038 standard; protein; 459 AA.
 XX AC AAW54038;
 XX DT 06-AUG-1998 (first entry)
 XX DE Human FK506 binding protein, FKBP52.
 XX KW FK506 binding protein; FKBP52; human; immune response regulator;
 XX KW immunosuppressant; steroid hormone receptor transformation.
 XX OS Homo sapiens.
 XX PN US5763590-A.
 XX PD 09-JUN-1998.
 XX PF 09-NOV-1994; 94US-00336618.
 XX PR 11-OCT-1991; 91US-00777752.
 XX PR 16-OCT-1992; 92US-00863325.
 XX PR 29-MAR-1994; 94US-00218989.
 XX PA (VERT-) VERTEX PHARM INC.
 XX PI Peattie DA, Livingston DJ, Harding MW;
 XX WPI; 1998-347419/30.
 XX N-PSDB; AAV24010.
 XX DNA sequence encoding human FK506-binding protein - and recombinant DNA molecule containing it.
 XX Example 5; Fig 3; 28pp; English.

This sequence is 52 kD human FK506 binding protein, referred to as FKBP52, of the invention. The FKBP52 protein plays a key role in regulating immune responses. FKBP52 may be useful for mediating steroid hormone receptor transformation. The DNA may be used to screen for new immunosuppressants, and in assays for metabolites in samples from

CC individuals taking immunosuppressants. The DNA may also be used in assays
CC for identifying natural intracellular rapamycin-like or FK506 like
CC substances, and in assays for identifying natural intracellular
CC substrates that are potential targets for other immunosuppressants
XX
SQ Sequence 459 AA;

Query Match 8.3%; Score 88.5; DB 2; Length 459;
Best Local Similarity 24.6%; Pred. No. 1.8;
Matches 30; Conservative 25; Mismatches 50; Indels 17; Gaps 3;
QY 18 DPHYDKITEINKAIDDAIAIEQSETI-----DPMKVPDHD-KFERHV 62
DB 75 DKSFGLGKEVIAKWDIAIAIMKVGCHITCKPEYVAGSAGSPKIPPNATLVFE--V 132
QY 63 GIVDFKGLAMRNIEARGLKQKQGDANVKGEIGVKAHLIGVHDDIVSMEYDLAYKL 122
DB 133 ELFEFKGEDLTEEDGGIIRRIQTRGEGYAKPNEGAIVEVALEGGYKDKLPDQRELRFEI 192
QY 123 GD 124
DB 193 GE 194

RESULT 39
ABR47454
ID ABR47454 standard; protein; 459 AA.
AC ABR47454;
XX
XX 12-JUN-2003 (first entry)
DE Breast cancer associated protein sequence SEQ ID NO:140.
DE Human; breast cancer; cytostatic; gene therapy.
KW Homo sapiens.
OS
XX WO2003004989-A2.
XX 16-JAN-2003.
XX 21-JUN-2002; 2002WO-US019669.
XX 21-JUN-2001; 2001US-0299887P.
PR 27-JUN-2001; 2001US-0301572P.
PR 18-JUL-2001; 2001US-0306501P.
PR 25-SEP-2001; 2001US-0325002P.
PR 05-MAR-2002; 2002US-0362585P.
PR 14-MAY-2002; 2002US-0380391P.
XX (MILL-) MILLENIUM PHARM INC.
XX Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamatkar S;
XX Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;
XX Bast RC, Hortobagyi GN, Puzstai L, Meric F, Sahin A, Mills GB;
XX WPI; 2003-210381/20.
XX N-ESDB; ACC50146.

Breast cancer diagnosis or treatment by comparing the level of expression
of a marker in a patient sample with that in the control non-breast
cancer sample.
XX Claim 1; SEQ ID NO 140; 128pp; English.
XX The present invention describes a method for assessing whether a patient
is afflicted with breast cancer. The method comprises comparing the level
of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
ABR47386 to ABR47632) in a patient sample and the normal level of
expression of the marker in a control non-breast cancer sample, where a
significant increase in the level of expression of the marker in the
patient sample and the normal level is an indication that the patient is

CC afflicted with breast cancer. The breast cancer associated sequences from
CC the present invention have cytostatic activities and can be used in gene
CC therapy. The method is useful for diagnosing and treating breast cancer.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 459 AA;

Query Match 8.3%; Score 88.5; DB 6; Length 459;
Best Local Similarity 24.6%; Pred. No. 1.8;
Matches 30; Conservative 25; Mismatches 50; Indels 17; Gaps 3;
QY 18 DPHYDKITEINKAIDDAIAIEQSETI-----DPMKVPDHD-KFERHV 62
DB 75 DKSFGLGKEVIAKWDIAIAIMKVGCHITCKPEYVAGSAGSPKIPPNATLVFE--V 132
QY 63 GIVDFKGLAMRNIEARGLKQKQGDANVKGEIGVKAHLIGVHDDIVSMEYDLAYKL 122
DB 133 ELFEFKGEDLTEEDGGIIRRIQTRGEGYAKPNEGAIVEVALEGGYKDKLPDQRELRFEI 192
QY 123 GD 124
DB 193 GE 194

RESULT 40
ABG72858
ID ABG72858 standard; protein; 459 AA.
XX ABG72858;
AC ABG72858;
XX 25-FEB-2003 (first entry)
DE Monoconal antibody preparation method related protein #3.
DE Transformation; antibody; peptidyl prolyl isomerase; pPiase;
KW chaperone-like activity; monoclonal antibody preparation; human.
XX Homo sapiens.
OS
XX JP2002262883-A.
XX 17-SEP-2002.
XX 13-MAR-2001; 2001JP-00070928.
XX 13-MAR-2001; 2001JP-00070928.
XX (SEKI) SEKISUI CHEM IND CO LTD.
XX (KAY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.
XX WPI; 2003-132125/13.
XX N-PSDB; ABX13847.

A transformant useful for preparing a monoclonal antibody comprises an
antibody gene and a gene encoding pPiase.
XX Disclosure; Page 12-13; 16pp; Japanese.
XX The invention describes a transformant containing an antibody gene and a
gene encoding a peptidyl prolyl isomerase (pPiase) having chaperone-like
activity. The methods detailed using the transformant are useful for the
preparation of a monoclonal antibody. This is the amino acid sequence of
a protein associated with the method of preparing a monoclonal antibody
described in the invention
XX
SQ Sequence 459 AA;

Query Match 8.3%; Score 88.5; DB 6; Length 459;
Best Local Similarity 24.6%; Pred. No. 1.8;
Matches 30; Conservative 25; Mismatches 50; Indels 17; Gaps 3;

Qy	30	NKAIODATAAQESQETIDPMKKVPDPAOKFBRHVGIYDFKGELAMRNIEARGUKQMQRQSD	89
Dd	21	NRIAGERISIVEDVEGVTRDIYATGEWLNRFSMTDTGG---IDDVDAPFMEEQIKHQAE	77
Qy	90	AN-----VKRGEIGV-K-----AHLIIGVHDDIV-----SMEYDL---AYK	121
Dd	78	TAMEEADUVIFVWSGKGSGITDADEYVARKLYTKHPVILAVNKVDNPEMRNDIFYDALG	137
Qy	122	JGLDHPTTHV----ISDIQPFWALSLEISDEGNITMTSPVRQFANV-----VNHGGL	172
Dd	138	LGEPIPISSVHGICTGDVLVAIVENLNEYEEENPDVIKSLICGRPNVGKSSLIINALICE	197
Qy	173	S--ILDPIFGVLSVLTALFODTVVRKEMTKV	201
Dd	198	DRVTASPVAGTTRDAIDTHFTDTDGGEFTMI	228

RESULT 42
AAB04108
ID AAB04108 standard; protein; 436 AA.

Streptococcus pneumoniae protein sequence ID36.

DE yphC protein of *Streptococcus pneumoniae* (GTP binding protein).

KW pleural empyema; endocarditis.
 KW otitis media; conjunctivitis; pneumoniae; meningitis; sinusitis;
 KW GIP binding protein; antibody; treatment; bacteremia
 KW ypnC; GIP binding protein; antibody; treatment; bacteremia
 KW pleural empyema; endocarditis.

PD 10-FEB-2000.

27-JUL-1998; 98GB-00016336.

XX PA (MICR-) MICROBIAL TECHNICS LTD.

XX
DR
WPT: 2000-195301/17

Streptococcal proteins and polynucleotides useful for diagnosis, PT treatment and prophylaxis of bacterial infections.

This sequence represents a *Streptococcus pneumoniae* protein of the invention. The proteins (or their homologues, derivatives and/or fragments) are useful as immunogens or antigens. Immunogenic or antigenic compositions comprising the proteins are useful as vaccines and also in diagnostic assays. The sequences are useful for the detection or diagnosis of *S. pneumoniae* infection, by contacting a sample to be tested with them. Agents capable of antagonising, inhibiting or interfering with the function or expression of the protein or polypeptide are useful in medical compositions in the treatment or prophylaxis of *S. pneumoniae* infection. As the sequences can be used to treat *S. pneumoniae* infection, they can be used to treat bacterial pneumonia, which has high rates in young children, the elderly, and in patients with predisposing conditions such as asplenia, heart, lung and kidney disease, diabetes, alcoholism, or with immunosuppressive disorders, especially AIDS. They can also be used to treat pneumococcal septicaemia, otitis media, sinusitis, and meningitis.

Query Match 8.2%; Score 88; DB 3; Length 436;
Best Local Similarity 21.8%; Pred. No. 1.9;
Matches 46; Conservative 34; Mismatches 89; Indels

30 NKAIDDAIAAEIQSETIDPMKVPDHADKFRHHVGI VDFKGELAMRNIEARGGLKQMKRQGD 89

Db 21 NRIAGERISIVEDVEGVTDRIVATGEWLNRFSMIDTGG---IDDDVDAFPMEQIKHQA 77
 Qy 90 AN-----VKGEIGIVK-----AHLIGVHDDIV-----SMEYDL-----AYK 121
 Db 78 IAMEEADVIVFVSGKEGIDTADYVARKLYKTHKPVILAVNKVDNPEMNDIYDFYALG 137
 Qy 122 LGDLHPHTHV-----ISDIQDFVVALSLEISDEGNITMTSFEVROFANV-----VNHIIGL 172
 Db 138 LGEPLPISVHGIGTGVDLDAIVENLPNEYEEENPDVKFSLGRPNVKGKSLINAILGE 197
 Qy 173 S--ILDRIFGVLSDLTAIFODTVRKEMTKV 201
 Db 198 DRVIASPVAGTTRDAIDHTDIDTGDQEFMTI 228

RESULT 43
 ABU02185
 ID ABU02185 standard; protein; 436 AA.
 XX
 AC ABU02185;
 DT 23-OCT-2003 (revised)
 DT 11-FEB-2003 (first entry)
 XX
 DE S. pneumoniae type 4 strain protein from coding region #1762.
 XX
 KW Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
 KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;
 KW gene therapy; vaccine.
 XX
 OS Streptococcus pneumoniae; type 4 strain.
 XX
 PN WO200277021-A2.
 PD 03-OCT-2002.
 PF 27-MAR-2002; 2002WO-IB002163.
 PR 27-MAR-2001; 2001GB-00007658.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Masignani V, Tettelin H, Fraser C;
 XX
 DR WPI; 2003-040579/03.
 DR N-PSDB; ABX07474.
 XX
 PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,
 PT useful as medicaments for treating or preventing a disease or infection
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
 PT ear infection.
 XX
 PS Claim 1; SEQ ID NO 3524; 56pp; English.
 XX

The invention relates to a protein comprising or having at least 50% identity to any of the 2469 amino acid sequences, identified in the specification (available on a computer readable format), or its fragment, expressed from 2469 of 2489 identified DNA coding regions from the Streptococcus pneumoniae type 4 strain genomic sequence appearing as ABS56454. Also included are an antibody which binds one of the proteins, treating a patient by administering the protein, DNA or antibody (in a composition), a kit comprising first and second primers, which are the nucleic acid cited above or fragments between nucleotides 8-100 of a sequence not defined in the specification, for amplifying a target sequence contained within a Streptococcus nucleic acid sequence, where the first primer is substantially complementary to the target sequence and the second primer is substantially complementary to the complement of the target sequence, and where the parts of the primers having substantial complementarity define the termini of the target sequence to be amplified, assay comprising contacting a test compound with the protein, and determining whether the test compound binds to the protein and a Streptococcus pneumoniae bacterium, where one or more genes

CC encoding the proteins has been rendered inactive. The proteins, nucleic acid molecules, antibody and compositions are useful as medicaments for treating or preventing a disease or infection due to streptococcus bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis media or ear infection. They are also useful in developing vaccines, CC diagnostics and antibiotics. The methods are useful for identifying CC immunodominant proteins. The present sequence is one of the 2469 proteins expressed by the identified coding regions from the genomic sequence. CC Note: The sequence data for this patent did not form part of the printed CC specification, but was obtained in electronic format directly from WIPO CC at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to CC standardise OS field)
 XX
 SQ Sequence 436 AA;

Query Match 8.2%; Score 88; DB 6; Length 436;
 Best Local Similarity 21.8%; Pred. No. 1.9;
 Matches 46; Conservative 34; Mismatches 89; Indels 42; Gaps 8;
 QY 30 NKAIDDAIAAEQSTIDPMKVPDHADKFERHGVIVDFKGLAMRNIEARGLKQMKRGD 89
 Db 21 NRIAGERISIVEDVEGVTDRIVATGEWLNRFSMIDTGG---IDDDVDAFPMEQIKHQA 77
 QY 90 AN-----VKGEIGIVK-----AHLIGVHDDIV-----SMEYDL-----AYK 121
 Db 78 IAMEEADVIVFVSGKEGIDTADYVARKLYKTHKPVILAVNKVDNPEMNDIYDFYALG 137
 QY 122 LGDLHPHTHV-----ISDIQDFVVALSLEISDEGNITMTSFEVROFANV-----VNHIIGL 172
 Db 138 LGEPLPISVHGIGTGVDLDAIVENLPNEYEEENPDVKFSLGRPNVKGKSLINAILGE 197
 QY 173 S--ILDRIFGVLSDLTAIFODTVRKEMTKV 201
 Db 198 DRVIASPVAGTTRDAIDHTDIDTGDQEFMTI 228

RESULT 44
 ABU46202
 ID ABU46202 standard; protein; 436 AA.
 XX
 AC ABU46202;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #31729.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA50072.
 XX

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Db 149 REFQHLYEALRSEADWIVGINATRALTKYDAQLSLGRVQPTTIOVVARQOEINHPK 208
Qy 140 VALSLEISDE-GNIT-----MTSFEVRQFANVNVHIGG-----LSILDPIF 179
Db 209 AKKYITLSTEIGLTLFOLSTNKQHMKTEDATQIANEIKHVEGNVDSVEKKVKSHPKPLY 268
Qy 180 GVL-----SDVLTALFQDITVRKEMTKVL 202
Db 269 NLTDLQOEAVQRYKMGPKETINTIONLYERHKVLTYPRTSDNYLTDMDVTIKERLYALL 328
Qy 203 APAFKRELE 211
Db 329 ATDYKSOVK 337

RESULT 46
ABP27827
ID ABP27827 standard; protein; 436 AA.
XX AC ABP27827;
XX DT 02-JUL-2002 (first entry)
XX DE Streptococcus polypeptide SEQ ID NO 4830.
XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX KW group A streptococcus; Streptococcus pyogenes; antibacterial;
XX KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX OS Streptococcus pyogenes.
XX PN WO200234771-A2.
XX PD 02-MAY-2002.
XX PF 29-OCT-2001; 2001WO-GB004789.
XX PR 27-OCT-2000; 2000GB-00026333.
XX PR 24-NOV-2000; 2000GB-00028727.
XX PR 07-MAR-2001; 2001GB-00005640.
XX PA (CHIR-) CHIRON SPA.
XX PA (GENO-) INST GENOMIC RES.
XX PI Telford J, Masignani V, Margarit Y Rosi, Grandi G, Fraser C;
XX PI Tettelin H;
XX DR WPI; 2002-352536/38.
XX DR N-PSDB; ABN68458.
XX PT New Streptococcus protein for the treatment or prevention of infection or
XX PT disease caused by Streptococcus bacteria, such as meningitis, and for
XX PT detecting a compound that binds to the protein.
XX PS Claim 1; Page 3645; 4525pp; English.
XX CC The invention relates to a protein (ABP25413-ABP30895) from group B
XX CC Streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS
XX CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
XX CC the specification. The proteins have antibacterial and antiinflammatory
XX CC activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and
XX CC antibodies that bind (1) are used in the manufacture of medicaments for
XX CC the treatment or prevention of infection or disease caused by
XX CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
XX CC Nucleic acids encoding (1) are used to detect Streptococcus in a
XX CC biological sample. (1) is used to determine whether a compound binds to
XX CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
XX CC used as a vaccine or diagnostic composition. The disease caused by
XX CC Streptococcus that is prevented or treated may be meningitis. Nucleic
XX CC acid encoding (1) may be used to recombinantly produce (1) and may be
XX CC used in gene therapy. Antibodies to (1) are used for affinity
XX CC chromatography, immunoassays, and distinguishing/identifying
XX CC Streptococcus proteins

XX SQ Sequence 436 AA;
Query Match 8.1%; Score 87; DB 5; Length 436;
Best Local Similarity 22.2%; Pred. No. 2.4;
Matches 47; Conservative 41; Mismatches 80; Indels 44; Gaps 11;
Qy 30 NKAIDDAIAAIQSETIDPMKVPDHDKXFERHVGIVDFKGLAMENIARGLKQMKRQ- 88
Db 21 NRIAGERISIDVEGVTRDRIYATGEWLNRFSLDTGG---IDVDAPFMEQIKHQQA 77
Qy 89 ----DAN-----VKGEEGIVKAH-----LLIGVHD-DIVSMEYDL-----AYK 121
Db 78 IAMEEADVIVFVSCKEGVTDADDEVSKLYRTNTPVILAVNKVDNPERNDIYDFYSLG 137
Qy 122 LGDLHPTTHV-----ISDQDFVA-LSLEISDEGNITMTSPEVROFANV-----VNHIGG 171
Db 138 LGDPYPVSSVHGIGTGDVLDIAIVENLPVEEABE-NDIIRFSLIGRPNVKGSSLLNAILG 196
Qy 172 LS--ILDRIFGVLSVLTALTAIFQDITVRKEMTKV 201
Db 197 EDVIVASPVAGTTRDAIDTHFTDADGQEFMTI 228

RESULT 47
ABU46470
ID ABU46470 standard; protein; 436 AA.
XX AC ABU46470;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #31997.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Streptococcus pyogenes.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX DR N-PSDB; ACA50340.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 74394; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies on a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Query Match 8.1%; Score 87; DB 6; Length 436;
Best Local Similarity 22.2%; Pred. No. 2.4;
Matches 47; Conservative 41; Mismatches 80; Indels 44; Gaps 11;

30 NXAIDDAALQSSSTIDPMKVPDCHADKFERHVGIVDFKGLAMNIEARGLKQMKRQG- 88
21 NRIAGERSIVDEGVETDRIVATGEWLNQFSLIDGG---IDDVDAPMEQIKHQAO 77
89 -----DAN-----VKGEIGIVKXAH-----LLIGVHD-DIVSMEDYL-----AYK 121
78 TAMEEADVIVFVSGKEGVTDADEVSKILYRTNTPVILAVNKVDNPEMRNDIYDFVSLG 137
122 LGDLHPTTHV-----ISDIQDFWA-LSLEISDEGNITMTSFEVRQFANV-----VNHIGG 171
138 LGDPFVSVSHGIGTGDVLDALVENLPVEAAEE-NDDIIRSLIGRPNVGKSSLINAILG 196
172 LS--LIDPIFGVLSVLTAIFQDVTVRKEMTKVLA 201
197 EDRIASIPVAGTTRDAIDHTFTDADGQEFMTI 228

RESULT 48
ABBS4727
ID ABBS4727 standard; protein; 1640 AA.

ABBS4727;
29-AUG-2003 (revised)
16-MAY-2002 (first entry)
Lactococcus lactis protein p1314.
Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
Lactococcus lactis; IL1403.
FR2807446-A1.
12-OCT-2001.
11-APR-2000; 2000FR-00004630.
11-APR-2000; 2000FR-00004630.
(INRG) INRA INST NAT RECH AGRONOMIQUE.

Bolotine A, Sorokine A, Renault P, Ehrlich SD;
Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI; 2002-043418/06.
New nucleotide sequence useful in the identification of Lactococcus lactis and related species.
Claim 6; SEQ ID NO 1429; 2504pp; French.
The present invention is related to a Lactococcus lactis nucleotide sequence (ABA90521) and related proteins (ABBS3300-ABBS5621). The nucleic acid sequence is useful in the detection and/or amplification of nucleic acid sequence, particularly to identify Lactococcus lactis or related species. The proteins of the invention are useful for the biosynthesis or biodegradation of a composition of interest. The invention helps research in lactic bacteria, particularly useful in the production of yogurt and cheese. Note: The sequence data for this patent is based on equivalent patent WO2001/77334 (published 18-OCT-2001) which is available in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to standardise OS field)

Query Match 8.1%; Score 87; DB 5; Length 1640;
Best Local Similarity 20.7%; Pred. No. 16;
Matches 47; Conservative 45; Mismatches 73; Indels 62; Gaps 12;

16 SADPHYDKITSEINKAIDDAIALEQSETIDPMKVPDCHADKFERHVGIVDFKGLAMRN 75
1142 SMSKQYDEIVKNAQQRDDTISAARKQOT-----EVDKQAK--THDKTVELANSKADKN 1195
76 IEA-----RGLKQMKROGDANVKGEIGIVK-----AHLIT----- 105
1196 VKAAKEQGETVEQYTKGFKDSRLNLSNIDFGINGVLFNFKGNGNIGHVSLKGFATGTR 1255
106 GVHDDIVSVEYDLAYKGLDHPHTHVIS-----DIQDFVALSL---EISDEGNITMTS 156
1256 GLAQDFTALVGEFGELAH-HFSGIFAVGQGPBIRNLKAGTSILPHSMSKE-FLSLT- 1312
157 FEVRQFANVNHIGGLSILDPFGVLSVLTAIFQDVTVRKEMTKVLA 203
1313 -----ANLPAHADGVS-----GLSDALGWV--KSTYKDVTSVIS 1345

RESULT 49
ABU49623
ID ABU49623 standard; protein; 544 AA.

ABU49623;
19-JUN-2003 (first entry)
Protein encoded by Prokaryotic essential gene #35150.
Antisense; prokaryotic essential gene; cell proliferation; drug design.
Vibrio cholerae.
WO200277183-A2.
03-OCT-2002.
21-MAR-2002; 2002WO-US009107.
21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993.
25-OCT-2001; 2001US-0342923P.
08-FEB-2002; 2002US-00072851.
06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.
DR N-PSDB; ACA53493.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 77547; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 544 AA;
SQ
Query Match 8.1%; Score 86.5; DB 6; Length 544;
Best Local Similarity 24.4%; Pred. No. 3.7;
Matches 41; Conservative 21; Mismatches 69; Indels 37; Gaps 5;
QY 6 LIAAVAFVAVSA-----DPIHYDKITEINKAIDDAIAAEQSETIDPMKV 51
DB LAGGAVAVIKVGAATEVEMKEKKORVEDALHATRAAVEGVAGGVALLRAASKLSLV- 430
QY 52 PDHAKPERHVGIVDFKGLAMENIEARGLQKMKROGDA-----NVKGEIGVKAHLTI 105
DB ---GDNEEQNGVI-----RVLRAMEAPLQIVKNAGDEESVANNVRAGEGNYGNAAT 482
QY 106 GVHDDIVSMYDLAYKGLDHPHTHVLSIDIQDFWALSLEISBEIGNIT 153
DB 483 GYVGDMTEM-----GILDPTKVTRSAIQFAASVAGLMITTEAMIT 522
RESULT 50
AAU33974
ID AAU33974 standard; protein; 660 AA.
XX
XX AAU33974;
AC
DT 14-FEB-2002 (first entry)
XX
XX Staphylococcus aureus cellular proliferation protein #250.
DE
XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
KW

XX Staphylococcus aureus.
OS
XX WO200170955-A2.
PN
XX 27-SEP-2001.
PD
XX 21-MAR-2001; 2001WO-US009180.
PF
XX 21-MAR-2000; 2000US-0191078P.
PR
XX 23-MAY-2000; 2000US-0206848P.
PR
XX 26-MAY-2000; 2000US-0207727P.
PR
XX 23-OCT-2000; 2000US-0242578P.
PR
XX 27-NOV-2000; 2000US-0253625P.
PR
XX 22-DEC-2000; 2000US-0257931P.
PR
XX 16-FEB-2001; 2001US-0269308P.
PA
(ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
PI
XX WPI; 2001-611495/70.
DR N-PSDB; AAS51833.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Example 3; SEQ ID NO 5470; 511pp; English.
PS
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 660 AA;
SQ
Query Match 8.1%; Score 86.5; DB 4; Length 660;
Best Local Similarity 22.2%; Pred. No. 4.9;
Matches 36; Conservative 31; Mismatches 50; Indels 45; Gaps 5;
QY 1 MMKFLIIAAVAFVAVSADPIHYDKITEINKAIDDAIAAEQSETIDPMKV----- 51
DB 354 IFSLLMIALVSVAVAMFGNKEYETPDVIGKSVKEA-----EQIFKNKLGKLSRYSVD 409
QY 52 -----PDHAKPER--HVGIVDFK-----ELAMRNIEARGLQK 84
DB 410 KYPENEIKTPTGTERVERGSDVDVVISKGPVKMPVNIPLPKQALQKLSGLKDV 469
QY 85 KQOGDANVKGREGI-----VKAHLIIGVHDDIVSMEYDLAYK 121
DB 470 KIEKVYNNQAPKGYTANQSVTANTEIAIHDSNKLKLYESLGK 511
RESULT 51
AAU36586
ID AAU36586 standard; protein; 664 AA.
XX
XX AAU36586;
AC

XX 14-FEB-2002 (first entry)
XX Staphylococcus aureus cellular proliferation protein #756.
XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
XX Staphylococcus aureus.

OS WO200170955-A2.
PN 27-SEP-2001.
XX 21-MAR-2001; 2001WO-US009180.
XX 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX (ELIT-) ELITRA PHARM INC.
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
DR N-PSDB; AAS54445.
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX Example 3; SEQ ID NO 12179; 51pp; English.

XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 664 AA;
Query Match 8.1%; Score 86.5; DB 4; Length 664;
Best Local Similarity 22.2%; Pred. No. 4.9;
Matches 36; Conservative 31; Mismatches 50; Indels 45; Gaps 5;

QY 1 MKKELLIAAVFVAVSDPHYDKITEINKAIDATAAEQSETIDPMKV----- 51
DB 355 IFSLLMIALVSVFAMFGNKYEETPDVIGSVKEA-----EQIFNKNLKLGISRSYSD 410
QY 52 -----PDHAKFER--HVGIVDFK-----ELAMENIEARGLKQM 84
DB 411 KYPENEIKTTPNGVERGSDVDVVISKPEKVMNVGILPKEQALQKLSGLKDV 470
QY 85 KRQGANVKGEEI-----VKAHLIGVHDDIVSMYDLYAYKLGLHPTT--HVISDI 121
DB 471 KIEKYNNQAPKGYANGSVTANTEIAIHDSNKLVSGLTK 512

RESULT 52

ID ADB80059 standard; protein; 280 AA.
AC ADB80059;
XX 04-DEC-2003 (first entry)
XX Mycobacterium tuberculosis nutrient starvation-inducible protein #49.
DE mycobacteria; nutrient starving condition; mycobacterial latency;
XX mycobacterial infection; vaccine.
KW Mycobacterium tuberculosis.
OS WO2003004520-A2.
PN 16-JAN-2003.
PD 04-JUL-2002; 2002WO-GB003052.
PF 04-JUL-2001; 2001GB-00016385.
PR 05-OCT-2001; 2001GB-00023993.
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
PA James BW, Marsh P, Hampshire T;
XX WPI; 2003-210338/20.
DR N-PSDB; ADB80060.
XX New mycobacterial peptide, useful for the manufacture of a medicament for
PT treating or preventing, or a diagnostic reagent for identifying,
PT mycobacterial infection.

XX Claim 2; Page 208-209; 442pp; English.
XX The invention comprises the amino acid and coding sequences of
CC mycobacterial (Mycobacterium tuberculosis) proteins which are up-
CC regulated under nutrient starving conditions and maintain mycobacterial
CC latency. The DNA and protein sequences of the invention are useful for
CC the treatment, prevention and diagnosis of a mycobacterial infection. The
CC present amino acid sequence represents a Mycobacterium tuberculosis
CC protein of the invention.

XX Sequence 280 AA;

Query Match 8.1%; Score 86; DB 7; Length 280;
Best Local Similarity 28.3%; Pred. NO. 1.6;
Matches 34; Conservative 14; Mismatches 50; Indels 22; Gaps 3;

QY 33 IDATAAEQSETIDPMKVPHADKFERHVGIVDFKGLAMNIEARGLKQMKR----- 86
DB 128 LDPAAVADALIQATDLKVDVEPIERGRGKAAARALDIVDGAQSPKRWLL 187
QY 87 -----QGDANVKGEEIGVKAHLIGVHDDIVSMYDLYAYKLGLHPTT--HVISDI 135
DB 188 LIRAGFPRPQTQIAVRNENWGAELDVGWODIKVAEYD-----GDHILTSRYHKKDI 242

RESULT 53

ID AAG82219 standard; protein; 312 AA.
AC AAG82219;
XX 03-SEP-2001 (first entry)

XX S. epidermidis open reading frame protein sequence SEQ ID NO:1532.
XX Staphylococcus epidermidis SR1 strain; infection; diagnosis; vaccination;

QY 88 GDANVKG---EGIVKAHLI---GVHD--DIVSMEYDLAYKGLDLPHTTHVSDIQDFV 139
DB 1204 GDSVIKGOQLTDLGVPHHEILICGVRELQKLYLNEVQEVYRLQGV-----DINDKH 1255
QY 140 VAL-----SLEISDEGNITM---TSFEVROQFANV---VNHIG-----LSILDP 177
DB 1256 IEILVROMLQKVRITDPGDTLLFGEDVDKKEFEENRTEEDGKPAQAPVVLGITKA 1315
QY 178 IFGVLSVLTALFQDVTYRKEMTKVLAPA 205
DB 1316 SLGTESFISAASFQDT-----TRVLTD 1338

RESULT 55

AAU75885

ID AAU75885 standard; protein; 1441 AA.

XX AAU75885;

XX 08-MAY-2002 (first entry)

XX Human adhesion molecule protein AD3/BAA25490.1.

XX Human; adhesion molecule; AD3; BAA25490.1; cardiovascular disease;
XX atherosclerosis; ischaemia; thrombosis; haematological disease;
XX leukaemia; blood clotting disorder; cancer; brain tumour;
XX inflammatory disease; rhinitis; gastrointestinal disease;
XX Crohn's disease; respiratory disease; asthma; immune disorder;
XX rheumatoid arthritis; allergy; liver disease; cirrhosis; burn;
XX endocrine disease; diabetes; bone disease; osteoporosis; wound healing;
XX neurological disease; multiple sclerosis; bacterial infection;
XX Mycobacterium tuberculosis infection; viral infection.

XX Homo sapiens.

XX WO200208423-A2.

XX 31-JAN-2002.

XX 24-JUL-2001; 2001WO-GB003318.

XX 24-JUL-2000; 2000GB-00018126.

XX 17-OCT-2000; 2000GB-00025447.

XX (INPH-) INPHARMATICA LTD.

XX Phelps CB, Fagan RJ, Gutteridge A;

XX WPI; 2002-155219/20.

XX N-PSDB; ABK14973.

XX Five novel proteins, termed KIA0301, G7C, KIA0564, CAB01991.1 and
XX Rv0368c, which have been identified as adhesion molecules, useful in the
XX treatment and diagnosis of disease such as a cardiovascular disease,
XX cancer and immune disorders.

XX Claim 1; Fig 25; 302pp; English.

XX The invention relates to six novel proteins, termed KIA0301, G7C,
XX BAA25490.1, AAD21820.1, CAB01991.1 and CAA17374.1, which have been
XX identified as adhesion molecules (also termed AD1-6), their encoding
XX nucleic acids, fragments and functional equivalents. Also included are a
XX vector comprising the nucleic acids, a host cell transformed with the
XX vector, an AD-inhibitory ligand, a compound that either increases or
XX decreases the level of expression or activity of the AD protein, a
XX vaccine comprising the protein or nucleic acid, a method for the
XX identification of a compound that is effective in the treatment and/or
XX diagnosis of disease, comprising contacting AD or its nucleic acid with
XX one or more compounds suspected of possessing binding affinity for the
XX polypeptide or nucleic acid molecule, and selecting a compound that binds
XX to AD or its nucleic acid. The polypeptides, nucleic acids, vector,
XX ligand, or compound are useful in therapy or diagnosis of disease such as
XX a cardiovascular disease (including atherosclerosis, ischaemia,

CC restenosis, reperfusion injury, sepsis), a haematological disease (such
CC as leukaemia), a blood clotting disorder (such as thrombosis), cancer
CC (including lung, prostate, breast, colorectal and brain tumours,
CC metastasis), an inflammatory disease such as rhinitis, a gastrointestinal
CC disease (including inflammatory bowel disease, ulcerative colitis,
CC Crohn's disease), a respiratory disease (including asthma, chronic
CC obstructive pulmonary disease (COPD), respiratory distress syndrome,
CC pulmonary fibrosis), immune disorders (including autoimmune diseases,
CC rheumatoid arthritis, transplant rejection), allergy, liver diseases such
CC as cirrhosis, endocrine diseases such as diabetes, bone diseases such as
CC osteoporosis, neurological diseases (including stroke, multiple
CC sclerosis, spinal cord injury), burns and wound healing, bacterial
CC infection, particularly Mycobacterium tuberculosis infection, or virus
CC infection. The present sequence is the adhesion molecule protein, AD3
XX
XX

SQ Sequence 1441 AA;

Query Match 8.0%; Score 85.5; DB 5; Length 1441;

Best Local Similarity 22.0%; Pred. No. 19;

Matches 50; Conservative 35; Mismatches 81; Indels 61; Gaps 14;

QY 20 IHYXKI--TEEINKAIDDAIAIEQSEITDPMK-----VPHADKFERHV-----GIYDF 67

DB 343 LHRDTTVQTLTQPSVKDGLIVYEDSLPVKAVKLGHLVDEADKAPTNTCILKTLVE- 401

QY 68 KQELAM--RNTEARGLKQMKQGDANVKGEGIVKAH-----LLIGVHDDIVSMEYDLA 119

DB 402 NGEMILADGRRIVA-----NSANVNGRNVVVIHPDFRMIVLANRPGFPFLGNDF 452

QY 120 YKLGDLHPHTHVISD-----IQDFVVALS--LEISDEGNITWTS 156

DB 453 GTLGDIIF-SCHAVDNPKEHLEMLRQYGNVPPIQLKLVAAFGELSLADQGIINY-P 510

QY 157 FEVRQFANVNHIGLSILDPIFGVLSVLTALFQ-DTVRKEMTKVL 202

DB 511 YSTREVVNIHGLQKF-----PTEG-LSSVVRNVFDFSYNNDMREIL 552

RESULT 56

AAE10924

ID AAE10924 standard; protein; 2780 AA.

XX AAE10924;

XX 18-DEC-2001 (first entry)

XX Mouse monogenic audiogenic seizure-susceptible-1 (mass1) protein.

XX Mouse; monogenic audiogenic seizure-susceptible-1 gene; mass1;

XX transgenic animal; genetic abnormality; seizure.

XX Mus musculus.

XX Key Location/Qualifiers

XX Domain 24..58 "Conserved amino acid repeat motif of mass1

XX FT protein"

XX FT 150..184

XX FT /note="Conserved amino acid repeat motif of mass1

XX FT protein"

XX FT 277..311

XX FT /note="Conserved amino acid repeat motif of mass1

XX FT protein"

XX FT 535..569

XX FT /note="Conserved amino acid repeat motif of mass1

XX FT protein"

XX FT 650..684

XX FT /note="Conserved amino acid repeat motif of mass1

XX FT protein"

XX FT 788..802

XX FT /note="Conserved amino acid repeat motif of mass1

XX FT protein"

XX FT 882..916

[illegible]